MODELING OF MINORITY CARRIER RECOMBINATION AND RESISTIVITY IN SIGE BICMOS TECHNOLOGY FOR EXTREME ENVIRONMENT APPLICATIONS

A Thesis
Presented to
The Academic Faculty

by

Kurt A. Moen

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2008

MODELING OF MINORITY CARRIER RECOMBINATION AND RESISTIVITY IN SIGE BICMOS TECHNOLOGY FOR EXTREME ENVIRONMENT APPLICATIONS

Approved by:

Professor John D. Cressler, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor David S. Citrin
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Shyh-Chiang Shen
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: 12 November 2008

To my dear wife,

Jennifer

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to Dr. John D. Cressler for his guidance and support as my advisor. Both in the lab and in the classroom, his enthusiasm for research and teaching, combined with his sincere care for others, has shown me how to excel as an engineer while being a blessing to those around you.

I would like to thank the other members of my thesis advisory committee, Dr. David Citrin and Dr. Shyh-Chiang Shen. I would also like to sincerely thank all of the members of the SiGe Devices and Circuits research team at Georgia Tech for their support and collaboration: Dr. Lance Kuo, Dr. Bongim Jun, Stan Philips, Steven Finn, Anuj Madan, Jiahui Yuan, Laleh Najafizadeh, Ryan Diestelhorst, Marco Bellini, Aravind Appaswamy, Akil Sutton, Ted Wilcox, Curtis Grens, Tushar Thrivikraman, Tom Cheng, Prabir Saha, Nand Jha, Steven Horst, Partha Chakraborty, Duane Howard, Dylan Thomas, Sachin Seth, and others. I am grateful to Dr. Guofu Niu from Auburn University for helpful discussions regarding my research.

I am also thankful for the support of the NASA ETDP and NASA DTRA programs from which this research was primarily funded, the SiGe BiCMOS team at IBM Microelectronics for providing some of the hardware used in this research, and the Georgia Electronic Design Center at Georgia Tech for logistical support.

I am deeply indebted to my parents, Dale and Marghy Moen, for their love and care throughout my life. I am very grateful to my brother and his wife, Brent and Amy, for the sincere care and support that they have shown me. And most of all, I am beyond grateful to my wife Jennifer, who has made me a better person in so many ways, and who I have grown to love so dearly. Her constant support and care has sustained me over the course of this degree.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This work presents a summary of experimental data and theoretical models that characterize the temperature-dependent behavior of key carrier-transport parameters in silicon down to cryogenic temperatures.

In Chapter 1, the Silicon Germanium (SiGe) Heterojunction Bipolar Transistor (HBT) is reviewed, along with the catalysts for the increasing use of SiGe BiCMOS technology as an extreme-environment electronics platform. Subsequently, the low-temperature operation and radiation response of the SiGe HBT is summarized.

Chapter 2 presents an overview of minority-carrier recombination lifetime in silicon. SRH recombination theory is discussed, including the dependence of recombination lifetime on temperature and injection level. Carrier lifetime measurement methods are reviewed, and experiments to study carrier lifetimes in the substrate of a commercial SiGe BiCMOS process are presented. The experimental data is analyzed and calibrated TCAD-relevant models are developed.

Chapter 3 presents an overview of low-temperature resistivity in silicon. Modeling of resistivity over temperature is discussed, addressing the prevailing theoretical models for both carrier mobility and incomplete ionization. Experimental measurements of the temperature dependence of resistivity in both p-type and n-type silicon are presented, and calibrated TCAD-relevant models are developed.

In Chapter 4, a summary is given of the advances presented by this thesis. The ability to integrate the calibrated models presented in Chapters 2 and 3 within commercial TCAD software is demonstrated. In addition, applications for accurate cryogenic models of carrier-transport parameters are discussed, and future directions are outlined for research into cryogenic modeling of fundamental physical parameters.

CHAPTER I

INTRODUCTION

The primary incentives for developing a more accurate understanding of charge carrier dynamics are presented in this chapter. The fundamentals of the SiGe HBT and SiGe BiCMOS technology are described, along with the challenges posed to electronic systems by extreme environments, including both temperature and radiation effects.

1.1 Motivation

In the past decade, there has been extraordinary growth in the global telecommunications market, largely driven by key emerging broadband communications applications, such as mobile communications (GSM/CDMA), WLAN, GPS, DSL, and satellite communications. At the core of this growth is the development of monolithic integrated circuit (IC) technologies that have made it possible to build complex integrated systems at reasonable cost. GaAs technologies dominated the market at first, as the need for their high performance outweighed their relatively high cost. Recently, however, as silicon-based technologies have made significant performance gains, they have established a significant presence in many radio-frequency (RF), microwave, and millimeter-wave applications, due to their relatively low cost and capability of single-chip integration, which further reduces system cost and size [11].

Among the building blocks for broadband technologies are low-noise amplifiers, power amplifiers, and voltage-controlled oscillators. The key performance metrics for these blocks include low power consumption, high gain, high frequency operation, high dynamic range, good linearity, low noise. However, these circuit-level metrics are necessarily coupled to corresponding device level metrics; consequently, understanding device-level performance is critical to achieving circuit and system-level performance

gains. With this in mind, any useful analysis of device performance depends on a solid understanding of the fundamental physical processes that operate within the device.

One increasingly important segment of the broadband communications market is high-speed communication links with satellites in orbit around the earth. Extra-terrestrial electronics systems are required to operate in extremely harsh environments and are subjected to both particle radiation and cryogenic temperatures. Cryogenic temperatures induce significant changes in all aspects of device operation; thus, the fundamental physical properties that drive these changes must be accurately and thoroughly characterized in order to provide a solid foundation for the development of robust devices and circuits. Similarly, radiation effects introduce serious reliability concerns that must be addressed before a particular technology can become commercially viable for extreme-environment applications.

Radiation fields result from the magnetosphere and proton and electron belts surrounding the earth. Solar wind particles trapped in the earth's magnetic field result in the "Van Allen" radiation belts, which are particularly concerning for the orbital paths of satellites [28, 29]. As a result of operating in this extreme environment, electronic systems often suffer from degraded performance or altogether fail after a length of time. Device and circuit performance degradation can be attributed to three primary mechanisms: displacement damage, ionization damage, and single-event effects. This is of particular concern as technologies scale, since changes to the device structure and fabrication process can potentially lead to increased radiation sensitivity.

Over the past decade, silicon-germanium BiCMOS technology has emerged as a key enabling technology for RF and microwave circuit applications. While previously the use of III-V technologies was requisite for high-performance applications, the significant performance gains of the SiGe HBT vs. the Si BJT and its compatibility with standard Si processes have provided a technology path that provides a desirable combination of cost and performance. The SiGe HBT enables III-V competitive

performance for RF and microwave applications, while preserving the enormous yield, cost, and manufacturing advantages associated with conventional Si fabrication; having both the high speed HBT together on-chip with aggressively scaled CMOS allows great flexibility and increased integration in system design [10].

In the niche market of electronics for orbital and extra-orbital environments, SiGe BiCMOS technology promises to play an important role due to its unparalleled cryogenic temperature performance, radiation tolerance, wide temperature range capability, and optimal mixed-signal circuit design flexibility at the monolithic level. The technology is extremely robust to both displacement damage (DD) and total ionizing dose (TID) effects [4, 23, 42, 43]. Its demonstrated susceptibility to single event effects (SEE) [30, 38], however, must be addressed if SiGe BiCMOS is to be considered a legitimate space-qualified IC technology. Moreover, SEE must be considered in conjunction with cryogenic operation, since SEE sensitivity has been shown to increase at cryogenic temperatures [54]. At present, a wide variety of circuit architectures [30] and device-level techniques [53] have already been implemented in order to mitigate SEE with minimal impact on higher levels of system design. In order to further understand the nature of SEE sensitivity in SiGe digital ICs, the transient dynamics of charge generated by an heavy-ion strike must be thoroughly characterized.

The goal of this thesis is to develop a more accurate understanding of the physical phenomena that drive carrier transport in silicon across a wide temperature range. This analysis will be conducted within the context of a commercial SiGe BiCMOS technology, enabling many practical applications from the developments that will be presented. Resistance and carrier lifetime measurements are carried out across temperature and after proton irradiation. From the measured data, theoretical models are calibrated and adapted in order to accurately match the data across the entire temperature range. These models can then be applied towards a variety of predictive

TCAD simulations, such as simulations of the dynamic behavior of an HBT after being hit by an ion strike.

1.2 SiGe BiCMOS Technology

Although the first functional transistors ever demonstrated were fabricated from germanium, silicon quickly grew to dominate the industry due to its many commercial advantages, and Ge was more or less forgotten as a viable device material. Yet even early on the idea of combining silicon and germanium was suggested by Shockley as an interesting device concept, due their different band gaps and chemical compatibility. However, it would take over 30 years to develop the necessary sophistication in fabrication processes required to create device-quality SiGe films. This process culminated in the development of one of the first commercially viable SiGe HBTs by IBM in 1989 [35], using the relatively new ultra-high-vacuum / chemical-vapor-deposition (UHV/CVD) technique [31] to deposit a SiGe strained-layer as the base of the HBT. Since that time, the SiGe HBT has rapidly matured both in performance and commercial viability, and is currently offered by at least a dozen major semiconductor foundries [11].

The development of UHV/CVD was a critical step, because although Si/SiGe epitaxy has existed since the 1960's, it requires a high processing temperature that precludes its use in the manufacturing of high-speed devices, since any high temperature step would cause excessive out-diffusion of the thin base doping profile. Molecular beam epitaxy (MBE) is another longstanding technique that uses lower processing temperatures and produces reliable thin films [6]; however, despite the fact that it was used to fabricate the first functional SiGe HBTs in 1987 [15], it was not considered a manufacturable approach due to its incompatibility with patterned wafers. UHV/CVD, on the other hand, uses far lower processing temperatures (400-500 °C), is compatible with patterned wafers, and can be used to process batches of wafers rather than one at a time.

1.2.1 Strained Layer Epitaxy

Si and Ge are both Group IV elemental semiconductors with a diamond lattice structure. As such, the elements are entirely miscible regardless of the ratio of Si to Ge. Their respective lattice constants, however, differ significantly, placing constraints on both the fabrication process and corresponding device structures. For a given fraction of Ge, the SiGe alloy lattice constant is given by Vegard's rule,

$$a(Si_{1-x}Ge_x) = a_{Si} + x(a_{Ge} - a_{Si}) \tag{1.1}$$

where a is the lattice constant, and x is the Ge fraction. Experimental measurements of SiGe films show only minor departures from this linear dependence and can be fit by the following quadratic relationship [11]:

$$a(Si_{1-x}Ge_x) = 0.002733x^2 + 0.01992x + 0.5431(\text{nm}) \tag{1.2}$$

This lattice mismatch is a key concern when depositing a SiGe alloy onto a Si substrate with a smaller lattice constant. The deposited SiGe allow will distort in one of two ways. First, in what is known as "pseudomorphic" growth, the SiGe alloy will compress and adopt the underlying Si lattice constant, resulting in a perfect crystalline structure across the interface. This is naturally the most desirable result for device applications. On the other hand, the deposited alloy can "relax" during growth and revert to its natural lattice constant as determined by Equation 1.2. Relaxation occurs when misfit dislocations form in the film, breaking the crystallinity of the film across the interface and making the deposited layer unsuitable for device applications. This takes place when the strain energy in the film exceeds the activation energy required for misfit dislocation formation [10].

In practice, epitaxial SiGe films are deposited on top of a thin, undoped Si buffer layer, which ensures a pristine growth interface for the SiGe layer deposition. After the SiGe strained layer is deposited, a thin, undoped Si cap layer is deposited on

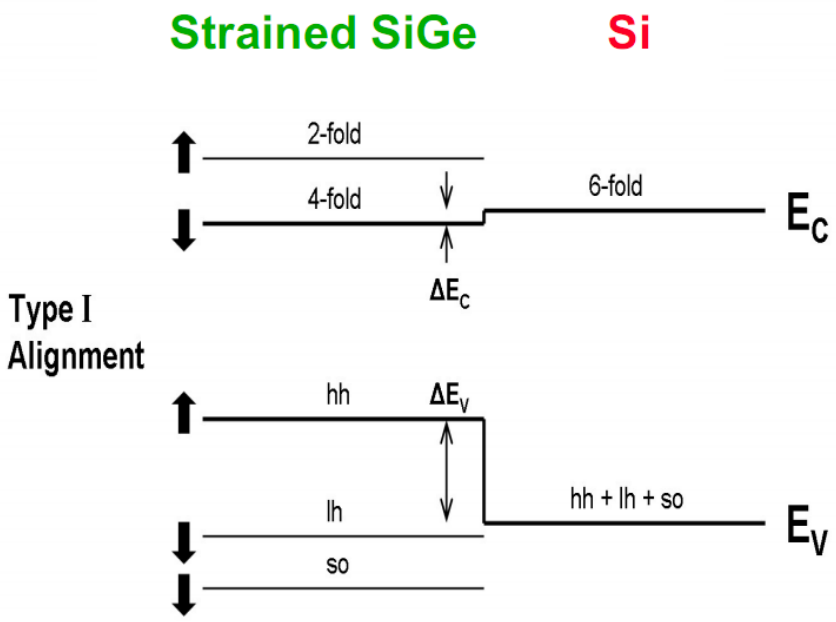


Figure 1.1: Schematic band alignments of strained SiGe grown on a Si substrate (After [11]).

top. This cap layer serves several purposes: it (1) improves SiGe film stability, (2) provides material for later oxidation steps, (3) provides room for the unavoidable out-diffusion of the boron base profile, and (4) can be used to introduce an active i-layer into the emitter-base junction that reduces the parasitic tunneling current. Theoretical approaches to film stability using force-balance theory have shown excellent agreement between theory and experiment for both CVD and MBE grown films when the influence of the Si cap layer is properly addressed [12]. With this knowledge, large-scale, high-yield production of devices incorporating reliable SiGe strained layers is feasible.

The resulting energy band structure of the SiGe alloy is critical to its utility as an active device layer. As demonstrated experimentally, a biaxially, compressively strained SiGe alloy has a smaller band gap than that of Si, a band offset that is predominantly in the valence band, and improved carrier transport parameters, all of which are desired properties. The band alignments of strained SiGe as grown on Si are

shown in Figure 1.1. In addition, there is substantial distortion of the density of states in the SiGe alloy, leading to a decrease of N_C and N_V as the Ge fraction increases. With regards to carrier transport parameters, there are varying degrees of agreement as to how Ge affects mobility. However, it is generally agreed that the hole mobility increases as the Ge fraction increases, while the electron mobility decreases [11].

1.2.2 SiGe HBT Operation

Fundamentally, the SiGe HBT is very much the same as its Si BJT counterpart, except that in the HBT a graded Ge profile is introduced in the base layer. The cross-section of a representative first generation SiGe HBT is shown in Figure 1.2. Its corresponding doping and Ge profiles are given in Figure 1.3. In the resulting energy band diagram (Figure 1.4), the Ge profile produces a graded offset that is primarily manifested in the conduction band. Although the inherent band offset due to the Ge profile occurs in the valence band, it is effectively translated to the conduction band. With a constant p-type doping in the base, both the Fermi level and the energy difference between the Fermi level and valence band are fixed; the Ge grading induces a valence band offset, but because the Fermi level must remain constant in equilibrium, it must decrease in

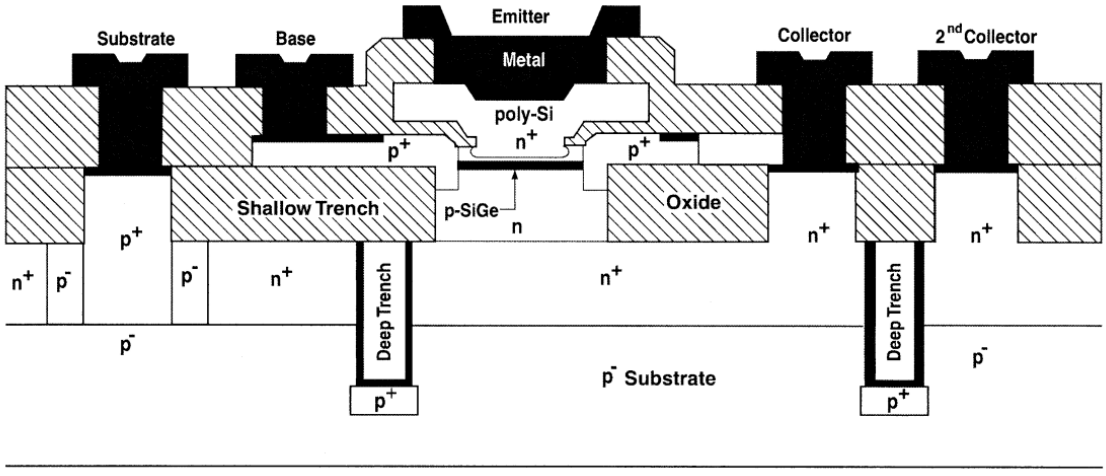


Figure 1.2: Cross-section of a representative first generation SiGe HBT (After [20]).

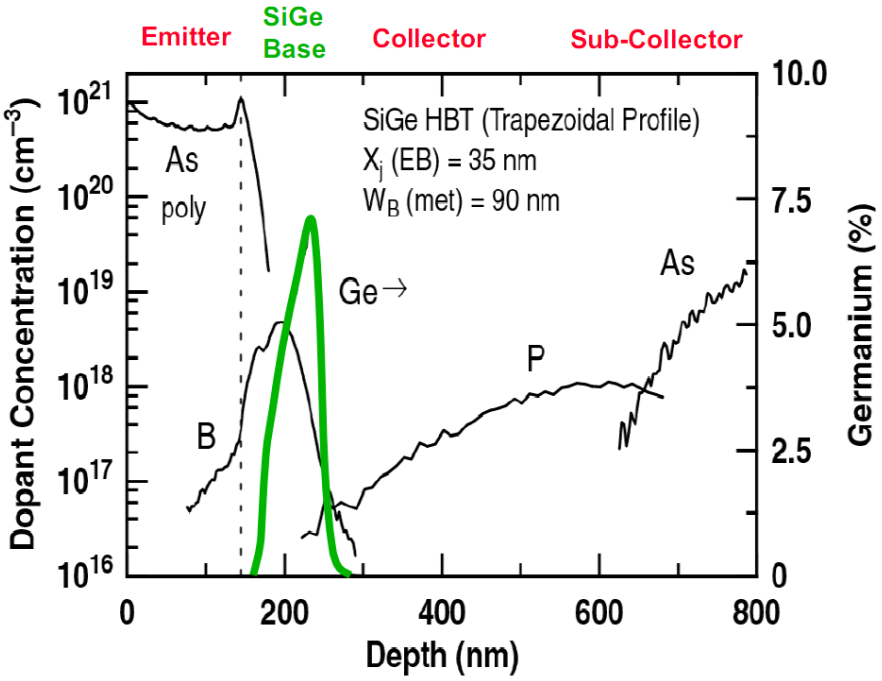


Figure 1.3: Measured SIMS profile of a representative first generation SiGe HBT (After [11]).

energy along with the conduction band edge.

For DC operation, one fundamental impact of the graded conduction band offset ($\Delta E_{g,Ge}(grade) = \Delta E_{g,Ge}(x = W_b) - \Delta E_{g,Ge}(x = 0)$) is to enhance minority electron transport across the base by inducing a drift field. In addition, the Ge content at the EB junction ($\Delta E_{g,Ge}(x = W_b)$) will reduce the potential barrier for electron injection from the emitter to the base, yielding exponentially greater electron injection for the same applied V_{BE}. Finally, a finite Ge content at the CB junction ($\Delta E_{g,Ge}(x = 0)$) will positively influence the output conductance of the transistor (ie. higher Early voltage), since the smaller base band gap near the CB junction effectively weights the base profile so that back side depletion of the neutral base with increasing V_{CB} is suppressed [10].

For AC operation, the Ge-grading induced drift field will intuitively lead to a reduced base transit time, which typically is the limiting transit time that determines

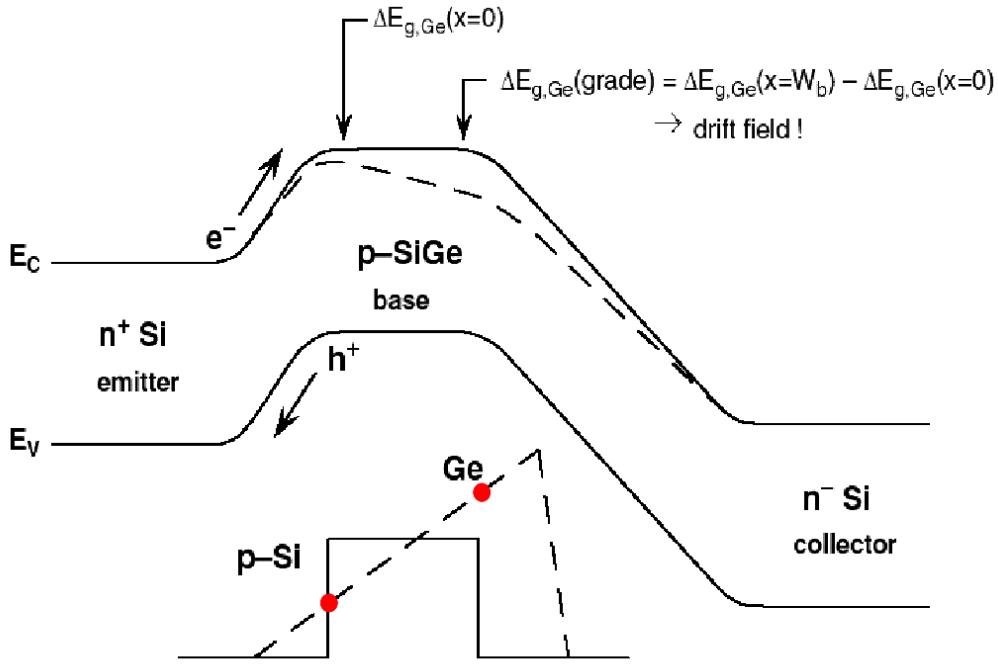


Figure 1.4: Energy band diagrams for a Si BJT and a graded-base SiGe HBT, biased in forward active mode at low-injection (After [11]).

performance metrics such as the maximum operating frequency. In addition, the Ge-enhanced injection of electrons from the emitter into the base dynamically produces a back-injection of holes from the base into the emitter. This reduces the emitter charge storage delay time, which is reciprocally related to the AC current gain of the transistor [10].

The theoretical consequences of the Ge-induced band gap changes to J_C can be derived by considering the generalized Moll-Ross collector current density relation,

$$J_C = \frac{q(e^{qV_{BE}/kT} - 1)}{\int_0^{W_b} \frac{p_b(x)dx}{D_{nb}(x)n_{ib}^2(x)}} \tag{1.3}$$

The presence of Ge is manifested within Equation 1.3 by both the carrier diffusivity D_{nb} and the intrinsic carrier density n_{ib}^2, given by Equation 1.4, with the position-dependent band gap narrowing described by Equation 1.5.

$$n_{ib}^2(x) = (N_C N_V)_{SiGe}(x)e^{-E_{gb}(x)/kT} \tag{1.4}$$

$$E_{gb}(x) = E_{gb0} - \Delta E_{gb}^{app} + [\Delta E_{g,Ge}(0) - \Delta E_{g,Ge}(W_b)]\frac{x}{W_b} - \Delta E_{g,Ge}(0) \tag{1.5}$$

Following the derivation in [10], we arrive at a formula for J_C, given by Equation 1.6. The parameters $\tilde{\gamma}$ and $\tilde{\eta}$ represent the effective density of states ratio and the minority carrier diffusivity ratio ($SiGe/Si$), respectively.

$$J_C = \frac{qD_{nb}}{N_{ab}^{-}W_b}(e^{qV_{BE}/kT} - 1)n_{io}^2 e^{\Delta E_{gb}^{app}/kT}\left\{\frac{\tilde{\gamma}\tilde{\eta}\dfrac{\Delta E_{g,Ge}(grade)}{kT}e^{\Delta E_{g,Ge}(0)/kT}}{1 - e^{-\Delta E_{g,Ge}(grade)/kT}}\right\} \tag{1.6}$$

From this expression, we can clearly see the collector current linearly increases with the Ge grading across the base, but exponentially increases with the amount of Ge-induced band gap narrowing at the EB junction. The J_C enhancement is evident when the standard Gummel characteristics for a comparably constructed Si BJT and SiGe HBT are plotted together (Figure 1.5). With the reasonable assumption that the

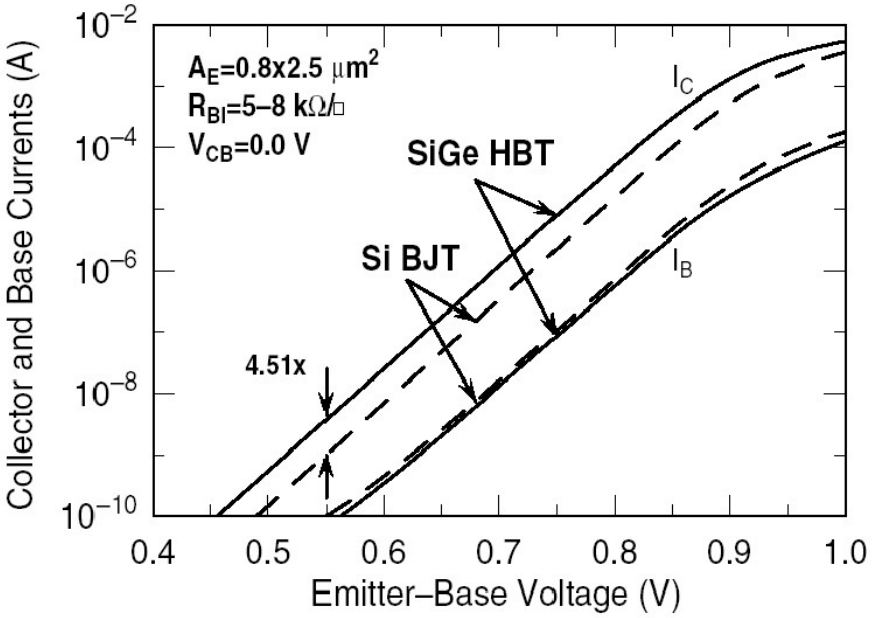


Figure 1.5: Comparison of current-voltage characteristics of a comparably constructed SiGe HBT and Si BJT (After [11]).

base currents are comparable, the current gain enhancement Ξ can be formalized as

$$\Xi = \left.\frac{\beta_{SiGe}}{\beta_{Si}}\right|_{V_{BE}} = \left\{ \frac{\tilde{\gamma}\tilde{\eta}\dfrac{\Delta E_{g,Ge}(grade)}{kT} e^{\Delta E_{g,Ge}(0)/kT}}{1 - e^{-\Delta E_{g,Ge}(grade)/kT}} \right\} \tag{1.7}$$

Due to the separate dependencies on the Ge grading and Ge content at the EB junction, it is important to note that the Ge profile can be tailored in order to achieve different levels of performance. This will be important when the effects of Ge on other important performance metrics are considered. Similarly to the current gain, the enhancement of the Early voltage can be quantified as

$$\left.\frac{V_{A,SiGe}}{V_{A,Si}}\right|_{V_{BE}} \simeq \left\{ \frac{e^{\Delta E_{g,Ge}(grade)/kT}}{\Delta E_{g,Ge}(grade)/kT} \right\} \tag{1.8}$$

In contrast to the current gain enhancement, the Early voltage enhancement is strictly dependent on the amount of Ge grading across the base. Since the total Ge content is limited due to stability constraints, a number of trade-offs are associated with the shape of the Ge profile; some performance metrics are most enhanced when there is a large Ge fraction near the EB junction (a "box" profile), while others are most enhanced with a maximal grading of Ge (a "triangle" profile).

The high-frequency performance of the device is largely determined by τ_b and τ_e, which are the base and emitter transit times, respectively. Following a similar process as before, we can derive an expression for the enhancement of τ_b due to the Ge-induced band gap narrowing [10],

$$\frac{\tau_{b,SiGe}}{\tau_{b,Si}} = \frac{2}{\tilde{\eta}} \frac{kT}{\Delta E_{g,Ge}(grade)} \left\{ 1 - \frac{kT}{\Delta E_{g,Ge}(grade)} \left[1 - e^{-\dfrac{\Delta E_{g,Ge}(grade)}{kT}} \right] \right\} \tag{1.9}$$

Likewise, the emitter transit time enhancement is derived to be

$$\frac{\tau_{e,SiGe}}{\tau_{e,Si}} = \frac{J_{C,Si}}{J_{C,SiGe}} = \left\{ \frac{1 - e^{-\Delta E_{g,Ge}(grade)/kT}}{\tilde{\gamma}\tilde{\eta}\dfrac{\Delta E_{g,Ge}(grade)}{kT} e^{\Delta E_{g,Ge}(0)/kT}} \right\} \tag{1.10}$$

From these expressions, it is clear that there is no Ge profile shape that will simultaneously optimize all of these parameters. β is enhanced by a large $\Delta E_{g,Ge}(x = 0)$, while

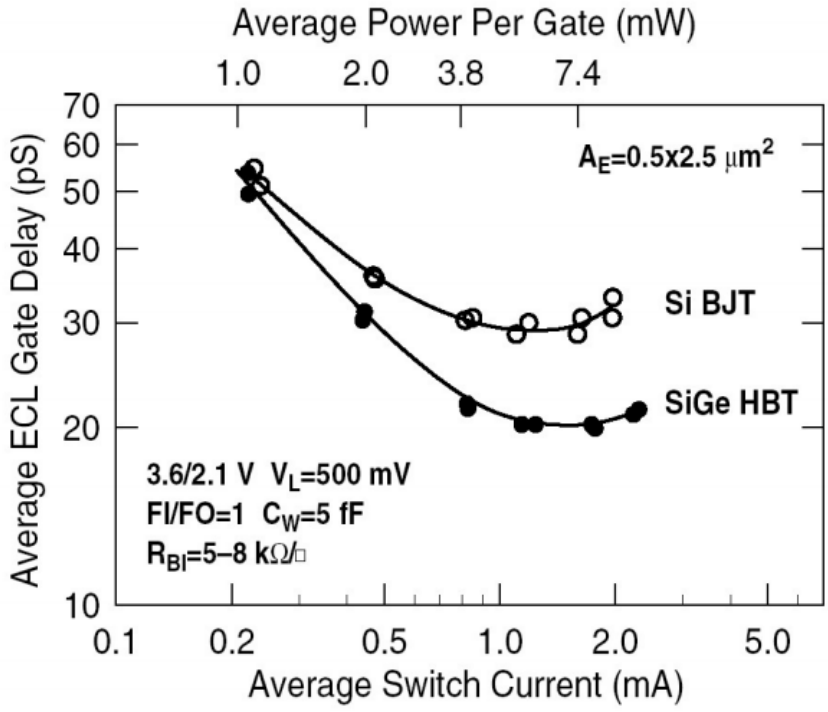


Figure 1.6: Comparison of ECL gate delay from comparably constructed SiGe HBTs and Si BJTs (After [11]).

a large $\Delta E_{g,Ge}(grade)$ enhances V_A, τ_b, and τ_e, which drives f_T, f_{MAX}, and the ECL gate delay. In practice, a compromise trapezoidal Ge profile shape is used, tailored to simultaneously enhance all performance metrics; for example, the same Ge profile is used to achieve the performance enhancements demonstrated in Figures 1.5 and 1.6.

1.3 Cryogenic Operation of SiGe Technology

The SiGe HBT operates very well at cryogenic temperatures, which is to be expected since band gap engineering generally has a positive influence on performance with decreasing temperature. Since it is a minority carrier device, the terminal currents are exponentially dependent on changes in the band gap, and as these changes are inevitably divided by the thermal energy (kT), they will be greatly magnified by any reduction in temperature. In contrast, the poor performance of Si BJTs at low temperatures has been long established, which has precluded its use for any cryogenic

electronics. In the SiGe HBT, then, the beneficial role of temperature can be leveraged in order to offset the inherent bandgap-narrowing-induced degradation in current gain that plagues the Si BJT. Consequently, reasonable or even exceptional cryogenic performance can be achieved by the SiGe HBT, as demonstrated by Figure 1.7, which shows the Ge-induced enhancement in current gain over temperature. Whereas the Si BJT experiences a degradation in current gain at lower temperatures, from Equation 1.7 we expect β_{SiGe} to increase exponentially with temperature [10].

Similarly, from Equation 1.8, we also expect the Early voltage to increase exponentially with temperature. Finally, the temperature dependence of the frequency response can be inferred from Equations 1.9 and 1.10. Both τ_b and τ_e should decrease with cooling; this is experimentally verified in Figure 1.8, which demonstrates increased f_T and decreased ECL gate delay for a SiGe HBT that is cooled to 85 K.

Combined with the benefits of low-power, scaled CMOS, the exceptional performance of the SiGe HBT across a wide temperature range makes the SiGe BiCMOS

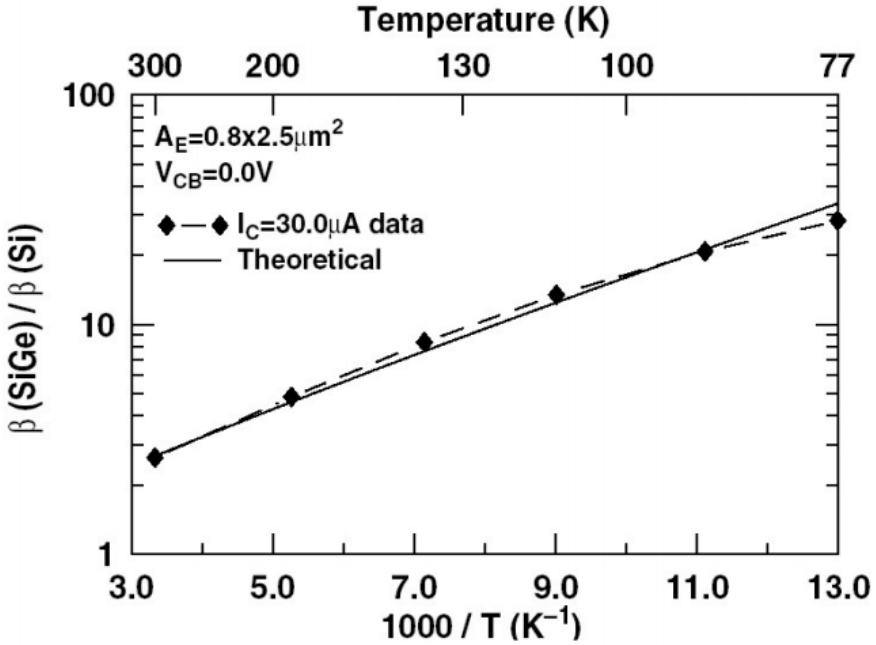


Figure 1.7: Current gain variation as a function of temperature (After [10]).

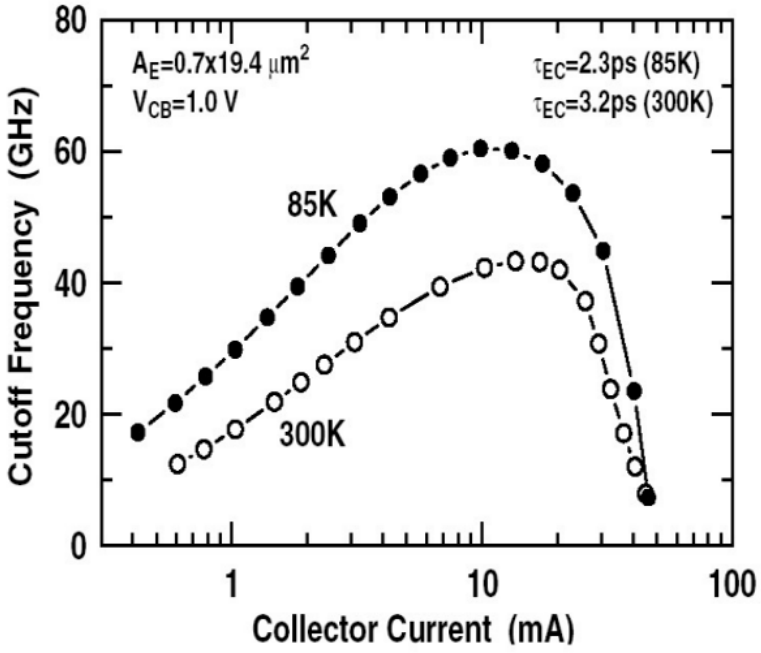


Figure 1.8: Cutoff frequency characteristics at 300 and 85 K for a SiGe HBT (After [10]).

platform an important contender for a number of niche cryogenic and extreme environment applications. The effects of radiation on the SiGe HBT will now be considered, since radiation hardness is an absolute necessity for any viable space-based electronics platform.

1.4 Radiation Effects

Microelectronic devices and circuits are used in a variety of radiation-intense environments, including satellites and other space-based electronics systems, radiological equipment for medical or security applications, control systems in nuclear power plants, semiconductor fabrication modules such as ion implantation and plasma etching, and high-energy particle physics experiments. In these environments, long-term device and circuit performance is gradually degraded by either displacement damage, in which the incident radiation displaces atoms from their site in the crystal lattice, or ionization damage, in which electron-hole pairs are generated within the device. The

severity of these types of damage, total ionizing dose (TID) and displacement damage (DD), is dependent on the total radiation dose that is absorbed. However, device and circuit reliability can also be affected by single-event effects (SEE), which are caused when charged particles strike a device, generating a funnel of charge along their path. SEE include both permanent errors such as single-event latchup (SEL), and transient errors such as single-event upset (SEU).

This thesis will primarily address the extra-terrestrial radiation environment, particularly the conditions to which orbital satellites are subjected. The types of particles found near the earth are (1) protons and electrons residing in the Van Allen belts, (2) heavy ions trapped in the magnetosphere, (3) protons and heavy ions from cosmic rays, and (4) protons and heavy ions from solar flares. Naturally, the entire radiation environment is heavily dependent on the solar cycle. A visual description of the radiation environment is given in Figure 1.9. The Van Allen belts consist primarily of low energy electrons (a few MeV) and high energy protons (several hundred MeV), along with a number of heavy ion particles trapped in the earth's magnetic field [51]. Galactic cosmic rays incorporate a range of elements with atomic numbers ranging from hydrogen to uranium, typically characterized by low flux levels and a very wide range of particle energies (a few MeV to several hundred GeV). Solar flares primarily introduce high energy protons and typically occur during periods of solar maxima [9].

As a result of all of these factors, electronic systems in orbit are subjected to particles and photons across a wide range of energies and flux. An ideal solution for space-based electronics would be a conventional IC technology that possesses all the necessary components for system-on-a-chip (SoC) integration, and is also inherently radiation-hard without any costly layout or process modifications. Since SiGe BiCMOS technology has demonstrated wide-temperature capability and has the potential for SoC integration, if it were also radiation-hard as fabricated, then it would present a very compelling solution as a space-based electronics platform.

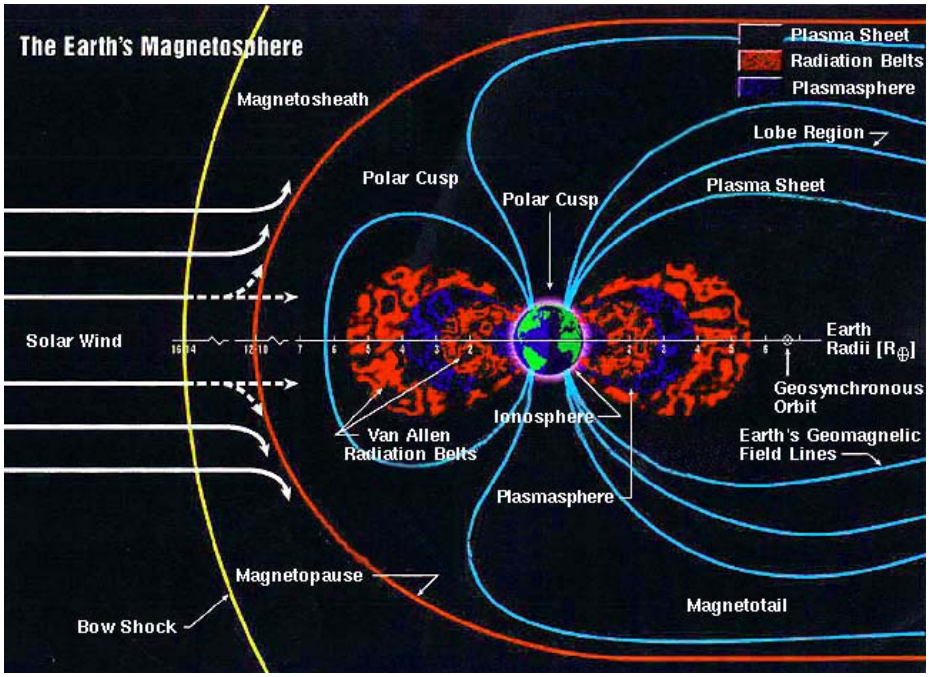


Figure 1.9: Solar wind and radiation belts surrounding the earth (After [33]).

1.4.1 Total Dose Effects

Displacement damage occurs when incident particles possess sufficient mass to displace individual atoms from their lattice sites, generating vacancy-interstitial pair defects. This can be caused by protons, neutrons, and heavy ions. A vacancy refers to the absence of an atom from its ordinary lattice site; an interstitial refers to an atom that is located in an off-grid position within the lattice. A vacancy-interstitial pair that is close together is referred to as a *Frenkel-pair* or a *close-pair*. Two adjacent vacancies are referred to as a *divacancy* and two adjacent interstitials are referred to as a *di-interstitial*. Additionally, impurities within the semiconductor can form defect-impurity complexes, such as the vacancy-oxygen complex or the interstitial-oxygen complex.

In silicon, high-energy neutron radiation can produce clusters of defects, whereas low-energy electron, gamma-ray, and proton radiation generally results in simple defects [22]. The presence of these defects is manifested by energy trap states at a wide

variety of positions within the band gap [8, 16]. These trap states can result in greatly increased Shockley-Read-Hall (SRH) generation and recombination rates. Methods by which these traps can be detected and characterized will be discussed at length in Chapter 2, including a detailed analysis of their influence on overall carrier lifetime. The tangible effects of traps on device operation are many: deep level traps tend to couple strongly to SRH recombination and generation of electron-hole pairs, whereas shallow traps result in electron or hole trapping, in which a carrier is temporarily trapped and later released. Traps at all energy levels can contribute to dopant or impurity compensation, possibly perturbing the majority carrier concentration and affecting parameters such as resistance and mobility. In addition, traps can lead to enhanced tunneling through potential barriers, which increases parasitic leakage currents. Finally, a more mechanical effect of defects is an increased lattice scattering of carriers, which results in lower carrier mobilities [22].

Experimental evaluations have shown SiGe HBTs to be extremely robust to DD and TID. As an example, consider the current gain response shown in Figure 1.10, taken from a fourth-generation SiGe HBT that was subjected to proton irradiation. Protons are an excellent test case, because they produce both displacement and ionization damage. After exposure to a sufficient flux of protons, the base current increases as expected due to increased recombination of carriers in the base. As described in [52], the excess base current is primarily associated with the EB spacer oxide at the periphery of the transistor; as such, ionization damage dominates the radiation response. The radiation hardness of the SiGe HBT is evidenced by the less than 30 % degradation in peak current gain up to a 1.0 Mrad (Si) dose. This result is mirrored across SiGe HBT technology generation [52].

Similarly, the AC performance of the SiGe HBT demonstrates robustness to TID. The cutoff frequency characteristics for first through fourth generation SiGe HBTs are shown in Figure 1.11. For first, second, and third generation devices, the

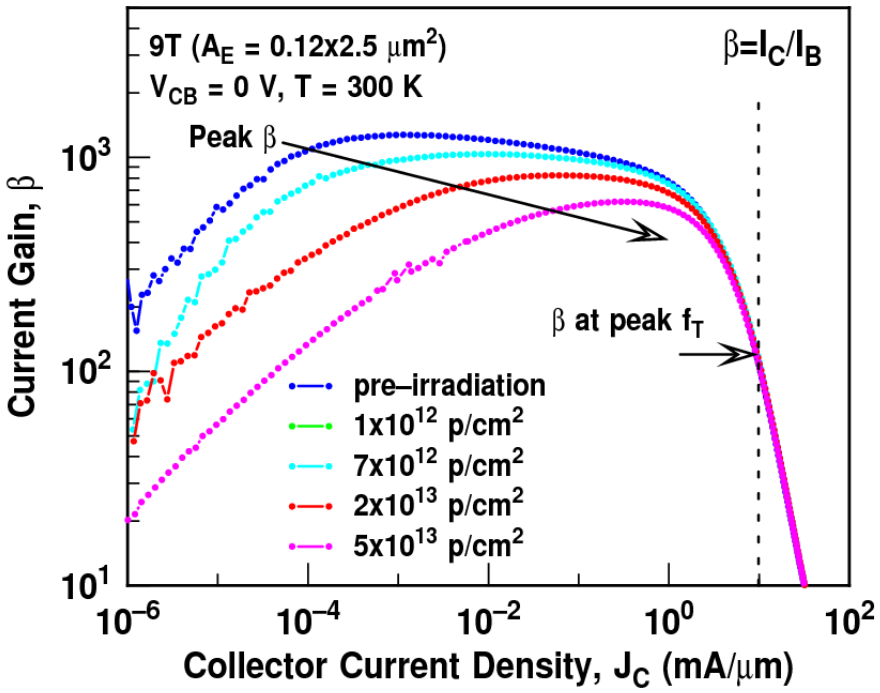


Figure 1.10: Radiation response of the current gain for a fourth-generation SiGe HBT (After [52]).

change in peak cutoff frequency is negligible. The fourth generation device exhibits a modest 6 % decrease that can be attributed to an increase in base resistance due to dopant deactivation and displacement damage [52]. Consequently, at typical operating currents, these results suggest that SiGe HBTs are inherently radiation hard to TID for typical orbital proton fluences.

1.4.2 Single Event Effects

The demonstrated susceptibility of SiGe HBTs digital logic circuits to single event effects (SEE) must be addressed if the SiGe BiCMOS platform is to be considered a legitimate space-qualified IC technology. This issue is even more critical since since SEE sensitivity has been shown to increase at cryogenic temperatures [54].

SEE can can be divided into two categories: (1) "soft" errors, which are transient errors that only temporarily affect the proper functioning of the device, and (2) "hard"

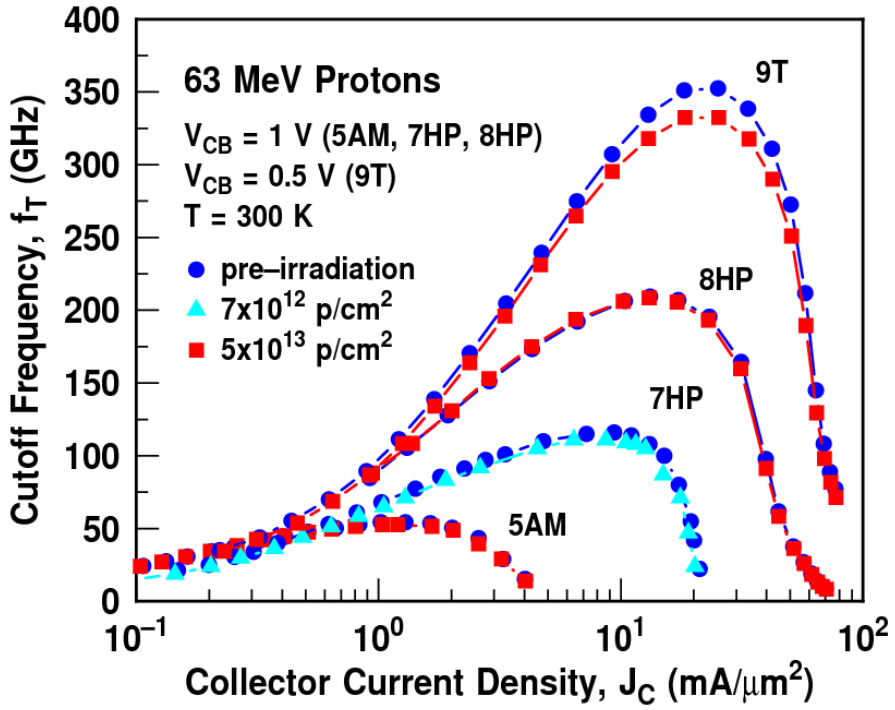


Figure 1.11: Radiation response of the cutoff frequency vs. bias current for first-, second-, third-, and fourth-generation SiGe HBTs (After [52]).

errors, which cause permanent changes or altogether destroy the device. Types of soft errors include:

- Single event upset (SEU) – a change of state or transient is induced by an ionized particle. This may occur in digital, analog, and optical components.
- Single event transient (SET) – a device error occurs that can be remedied by a reset or rewriting of the device.
- Multiple bit upset (MBU) – a single ionized particle causes multiple upsets or transients in a device.

Examples of hard errors include:

- Single event latchup (SEL) – an ion strike leads to the creation of a low impedance path from the supply voltage to ground. This can potentially lead to permanent device damage, but otherwise can be remedied by a power cycle of the device.

- Single hard error (SHE) – an SEU leads to a permanent change to the operation of the device.
- Single event burnout (SEB) – a high current state leads to the destruction of the device.
- Single event gate rupture (SEGR) – a single ion results in the formation of a conduction path through the gate oxide of a MOSFET.

The fundamental driving force behind the SEE response of a device is the amount of charge that is collected at each of its terminals following an ion strike. For the HBT, charge generated by an ion strike will ideally be collected by the substrate terminal rather than into an active node of the circuit Unfortunately, significant charge collection occurs at the active terminals of the device. The dynamics of charge collection and their manifestation within a circuit is heavily dependent on terminal bias, load condition, substrate doping, and ion strike depth, location, and angle. In order to understand SEE, calibrated 2D and 3D TCAD simulations of ion-strikes can be used to assess the charge collection characteristics of SiGe HBTs. This device-level knowledge can then be used to develop device-level mitigation schemes, as described in [53]. In addition, this device-level knowledge can be translated to circuit simulations of single event transients [34], allowing designers to develop circuit-level mitigation schemes; examples of this approach can be found in [10, 20, 32]. An example of the simulated SEU response of a SiGe digital logic circuit is given in Figure 1.12.

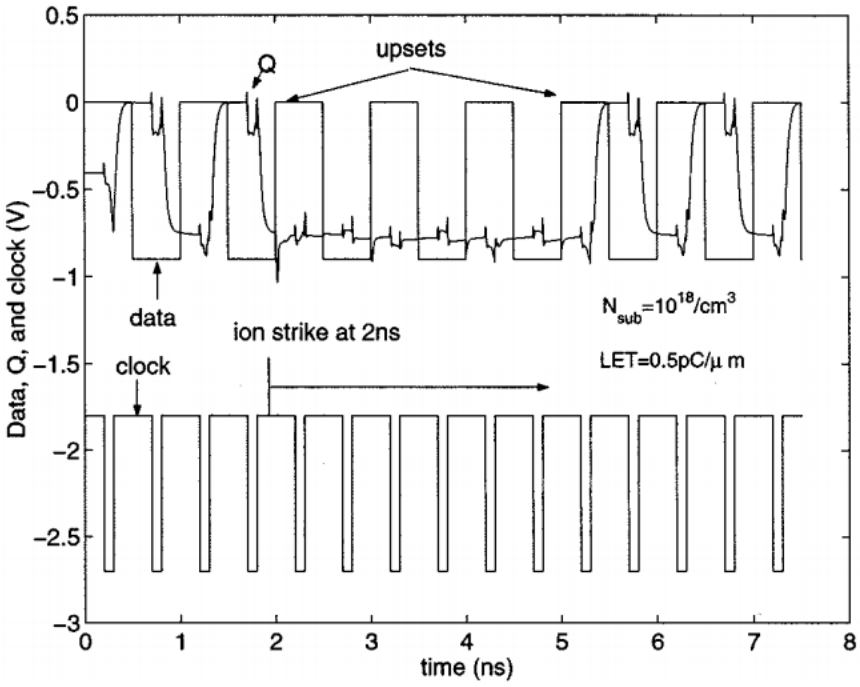


Figure 1.12: Simulated SEE response of a SiGe D-flip-flop (After [34]).

1.5 Summary

In this chapter, the catalysts for the increasing use of SiGe BiCMOS technology as an extreme-environment electronics platform have been outlined. The fabrication and structure of the SiGe HBT was reviewed and an overview was given of the DC and AC performance of the SiGe HBT. The challenges posed to electronic systems by extreme environments were addressed, including cryogenic temperatures and both total dose and single event radiation effects. Considering these challenges, a clear need has been established for developing a more accurate understanding of charge carrier dynamics within the SiGe HBT. With this in mind, the following two chapters will present an experimental and theoretical analysis of minority carrier lifetimes and resistivity in silicon.

CHAPTER II

MINORITY CARRIER LIFETIMES IN SILICON

In this chapter, an overview of recombination theory is given, with particular attention to Shockley-Read-Hall (SRH) recombination theory. The temperature and injection dependence of carrier recombination lifetimes will be addressed and measurement methods will be reviewed. Experimental data of the temperature and injection dependence of carrier lifetimes within a commercial SiGe BiCMOS technology are presented, including the effects of proton irradiation. The experimental data is analyzed and TCAD-relevant modeling is discussed.

2.1 Introduction

Generation in a semiconductor refers to processes by which electron-hole pairs are created. The energy required for an electron to be excited from the valence band to the conduction band can be supplied by thermal processes or by the absorption of photons. *Recombination*, on the other hand, refers to the inverse processes that lead to the annihilation of electron-hole pairs, in which an electron undergoes a spontaneous transition from the conduction band to an unoccupied state in the valence band. In order for energy and momentum to be conserved, the energy that is generated by these processes can either be released by photons or phonons or transferred to other carriers. The *recombination rate*, R, is directly proportional both to the relevant carrier densities and to a *recombination coefficient*, which reflects the quantum mechanical probability of an electron transition from the conduction band to the valence band; this probability is determined by the specific physical processes underlying recombination, which are referred to as *recombination mechanisms* [41].

In thermal equilibrium, the thermal generation rate, G_0, is equalized by an equilibrium recombination rate, R_0, leading to the equilibrium electron and hole densities, n_0 and p_0, whose product is constant, $n_0 p_0 = n_i^2$. When, for example, the generation rate is increased by an optical excitation, the recombination rate will also increase in order for reestablish a steady-state condition. In this state, the non-equilibrium electron and hole concentrations are increased, with $n_0 p_0 > n_i^2$. Subsequently, when the optical excitation ceases, thermal equilibrium cannot be reestablished instantaneously. The excess carrier densities, $\Delta n = n - n_0$ and $\Delta p = p - p_0$, decay along with the net recombination rate, $U = R - R_0$ [41]. Consequently, the net recombination rate is the parameter of interest, since it defines the dynamics of recombination for excess carriers. Assuming charge neutrality and $\Delta n = \Delta p$, the time dependent decay of the excess carrier density $\Delta n(t)$ is defined as

$$\frac{\partial \Delta n(t)}{\partial t} = -U(\Delta n(t), n_0, p_0) \tag{2.1}$$

In general, U can be written as a polynomial function of Δn, in which the zero-order coefficient must equal zero, since at thermal equilibrium $\Delta n = 0$ and $U(\Delta n = 0)$ must equal zero. Consequently, in the most basic case, U is directly proportional to Δn, and from Equation 2.1, the time dependent decay of Δn follows an exponential law. The time constant of this exponential decay is referred to as the *recombination lifetime*, also referred to as the *carrier lifetime*. This lifetime is defined generally as

$$\tau(\Delta n, n_0, p_0) := \frac{\Delta n}{U(\Delta n(t), n_0, p_0)} \tag{2.2}$$

When multiple recombination mechanisms are in effect, the total recombination rate is simply given as the sum of the individual recombination rates, and the effective carrier lifetime is given by the inverse sum of the reciprocal carrier lifetimes,

$$U_{eff} = \sum_i U_i \qquad \Longrightarrow \qquad \frac{1}{\tau_{eff}} = \sum_i \frac{1}{\tau_i} \tag{2.3}$$

From Equation 2.3, we can see that the effective lifetime is always smaller than the smallest contribution, and as a result the carrier lifetime is typically dominated by a

single recombination mechanism, as determined by the specific conditions within the semiconductor [41].

2.2 Recombination Mechanisms

The various recombination mechanisms within a semiconductor can be classified as either intrinsic or extrinsic. Intrinsic recombination mechanisms are always present, even in an ideal undisturbed crystal without impurities or defects. From the way the excess energy released by the annihilation of an electron-hole pair is dispersed, they can be distinguished as one of two distinct processes: (1) *radiative band-to-band recombination*, if the excess energy is released as a photon, or (2) *Auger band-to-band recombination*, if the excess energy is transferred to another carrier. For an indirect band gap semiconductor such as Si or SiGe, phonons must be involved in both of these processes, in order for momentum to be conserved. With extrinsic recombination mechanisms, the decay of an electron-hole pair does not occur directly via a band-to-band transition, but instead occurs stepwise via a defect center with an intermediate energy level within the band gap. This process is known as *Shockley-Read-Hall (SRH) recombination*, and although ideally it could be avoided by producing a perfect crystal, in practice it is often the dominant recombination mechanism. Imperfections in the crystal lattice and the unavoidable introduction of impurities during crystal growth and wafer processing lead to the presence of defect levels that drive SRH recombination. This recombination can occur either in the bulk of the crystal through lattice defects and impurities, or through surface states (*surface recombination*) typically caused by dangling bonds at the wafer surface and material interfaces [41].

Radiative band-to-band recombination is the direct annihilation of an electron-hole pair resulting in the emission of a photon with an energy approximately equal to that of the band gap. Since the process involves a conduction band electron moving to a vacant valence band state (hole), the radiative recombination rate is jointly dependent

on the concentrations of free electrons, n, and free holes, p. The following expression gives the net recombination rate U_{rad},

$$U_{rad} = B(np - n_i^2) \tag{2.4}$$

The coefficient B reflects the quantum-mechanical probability of a radiative transition, which is strongly dependent on the band structure of the crystal (i.e. whether it is direct or indirect). Assuming charge neutrality ($\Delta n = \Delta p$), when the non-equilibrium concentrations $n = n_0 + \Delta n$ and $p = p_0 + \Delta p$ are inserted into Equation 2.4, U_{rad} takes the form of a second-order polynomial; with the use of Equation 2.2, this results in the following expression for the radiative recombination lifetime τ_{rad},

$$U_{rad} = B(n_0 + p_0)\Delta n + B\Delta n^2 \qquad \Longrightarrow \qquad \tau_{rad} = \frac{1}{B(n_0 + p_0)\Delta n + B\Delta n^2} \tag{2.5}$$

While radiative recombination is typically the dominant recombination mechanism in direct semiconductors such as GaAs, since Si is an indirect semiconductor, it can typically be neglected [41].

Auger band-to-band recombination is traditionally viewed as a three-particle interaction in which the energy released by the recombination of an electron-hole pair is transferred to a third free carrier, which releases its excess energy as phonons to the crystal. If the third carrier is a conduction band electron, the Auger recombination rate U_{Auger} is proportional to n^2p; if the third carrier is a valence band hole, U_{Auger} is proportional to np^2. The net recombination rate is then given by

$$U_{Auger} = C_n(n^2p - n_0^2p_0) + C_p(np^2 - n_0p_0^2) \tag{2.6}$$

where C_n and C_p are the Auger coefficients for the third-carrier electron and third-carrier hole processes, respectively. Similar to Equation 2.5, a general expression for the Auger recombination lifetime can be derived. For the limiting case of low-level injection ($\Delta n \ll n_0 + p_0$), the relationships for n-type and p-type semiconductor are

$$\tau_{Auger}^{LLI,n} = \frac{1}{C_nN_D^2} \quad \text{for } n\text{-type}, \qquad \tau_{Auger}^{LLI,p} = \frac{1}{C_pN_A^2} \quad \text{for } p\text{-type} \tag{2.7}$$

Similarly, for high-injection ($\Delta n \gg n_0 + p_0$),

$$\tau_{Auger}^{HLI} = \frac{1}{(C_n + C_p)\Delta n^2} = \frac{1}{C_a \Delta n^2} \quad \text{for } n\text{-type and } p\text{-type} \tag{2.8}$$

where $C_a \equiv C_n + C_p$ represents the ambipolar Auger coefficient. From equations 2.7 and 2.8, the Auger lifetime is shown to inversely depend on the square of the carrier density. As a result, under low-level injection conditions, τ_{Auger} is injection-independent, while under high-level injection conditions, τ_{Auger} decreases as the injection density increases, its magnitude being independent of the doping concentration. Consequently, at high-injection, Auger recombination will dominate the net recombination rate over radiative recombination [41].

In reality, Auger recombination is much more complex than the idealized model presented above, and has been experimentally observed to be higher than predicted for doping levels below $1 \times 10^{18}\,\mathrm{cm}^{-3}$. The theory of a Coulomb-Enhanced (CE) Auger recombination has been proposed, in which the Coulomb interaction between the mobile charge carriers is taken into account. Due to this interaction, the electron density is locally increased in the vicinity of a hole and locally decreased in the vicinity of an electron. Since U_{Auger} strongly depends on the carrier concentration, local variations may change it significantly. Further discussion of this, including expressions for the CE-Auger recombination lifetime, can be found in [41].

SRH recombination is the dominant recombination mechanism in the lightly-doped bulk semiconductor, and as a result is the most important to understand when studying the transient behavior of charge deposited from an ion strike. The next two sections will present an review of SRH theory and an analysis of SRH lifetime behavior as a function of injection level and temperature.

2.3 SRH Recombination Theory

The presence of defects in a semiconductor, whether from impurities or from imperfections in the crystal lattice, produce discrete energy levels within the band gap.

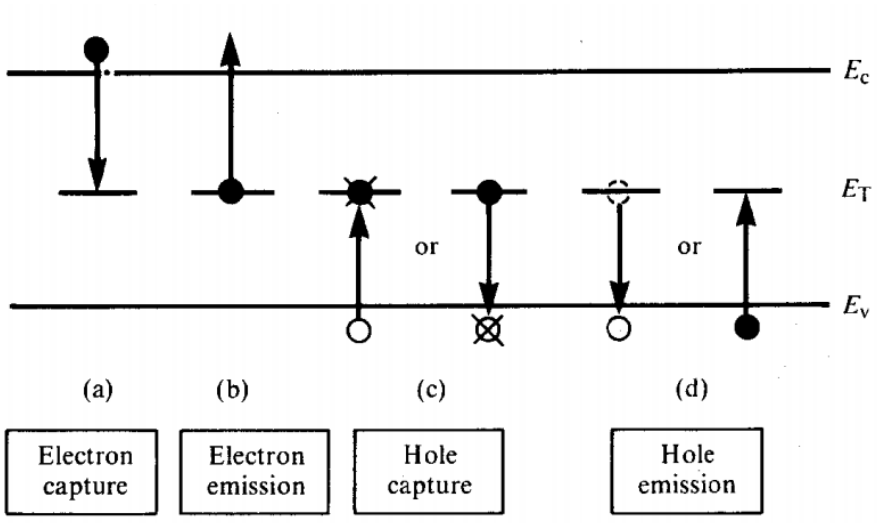


Figure 2.1: Energy band diagram showing the four elementary SRH processes (After [36]).

These defect levels may be filled with an electron or a hole, and may interact with the conduction or valence band through four elementary processes: (1) electron emission, (2) electron capture, (3) hole capture, and (4) hole emission. Figure 2.1 shows a diagram of these processes. From the two-step combinations of these processes, the defect level can act as a recombination center, generation center, or trap center.

As a recombination center, the following processes may occur: (1) a two-step capture process, in which a free electron from the conduction band and a free hole from the valence band are successively captured by the defect level, annihilating each other, or (2) a two-step relaxation process, in which a free electron from the conduction band relaxes to the defect level, then relaxes to the valence band, annihilating a hole.

As a generation center, the inverse of the recombination process occurs: (1) a free electron and free hole are successively emitted to the conduction and valence bands, respectively, or (2) an electron from the valence band is excited to the defect level and subsequently excited to the conduction band.

As a trap center, neither recombination nor generation occurs, and the defect center only interacts with one of the two energy bands. The trapping process consists of

either a hole being captured from and subsequently injected back to the valence band, or an electron being captured from and subsequently re-emitted to the conduction band [36].

The theory formulated by Shockley, Read, and Hall in 1952 stems from purely statistical considerations of the four elementary processes and establishes a recombination rate related to a single defect level as a function of the defect, material, and excitation parameters. In order to determine the time-dependent changes of the conduction band electron and valence band hole densities, the total emission and capture from all defects present must be determined; this begins by defining the emission and capture rates for a single defect center. The capture rates c_n^* and c_p^* may be expressed as

$$c_n^* = c_n n = (\sigma_n v_{th}) n \qquad \text{and} \qquad c_p^* = c_p p = (\sigma_p v_{th}) p \tag{2.9}$$

where $\sigma_{n,p}$ are the electron and hole capture cross-sections, respectively. To understand these expressions, consider the following illustration for the case of electrons: if movement within the lattice is considered to be relative to the position of a single electron, then the defect centers move randomly within the lattice at the thermal velocity v_{th} (1.0×10^7 cm/s). Consequently, each defect center, with a capture cross-section σ_n, clears a volume per unit time of $\sigma_n \times v_{th}$. The fact that any electrons within that volume will almost certainly be captured leads to Equation 2.9 [41].

Following the derivation in [41], the net time rate of change of the electron concentration is given by the electron emission rate minus the electron capture rate. Similarly, the net time rate of change of the hole concentration is given by the difference of the hole emission and capture rates. Equating the hole and electron rates of change, the net rate for SRH recombination through defects takes the form

$$U_{SRH} = \left(\frac{dn}{dt}\right)_{GR} = \left(\frac{dp}{dt}\right)_{GR} = \frac{(np - n_i^2)}{[N_t \sigma_n v_{th}]^{-1}(p + p_1) + [N_t \sigma_p v_{th}]^{-1}(n + n_1)} \tag{2.10}$$

where n_1 and p_1 are the SRH densities, named so because they equal the respective densities of electrons and holes when the Fermi level E_F coincides with the defect

energy E_T. These densities are defined as

$$n_1 := N_C \exp\left(-\frac{E_C - E_T}{kT}\right), \quad p_1 := N_V \exp\left(-\frac{E_T - E_V}{kT}\right) \tag{2.11}$$

The driving force of this recombination process is the term $(np - n_i^2)$, which describes the variation of the system from thermal equilibrium. By the same procedure followed to arrive at Equation 2.5, we can express the SRH lifetime as

$$\tau_{SRH} = \frac{\tau_{n0}(p_0 + p_1 + \Delta n) + \tau_{p0}(n_0 + n_1 + \Delta n)}{p_0 + n_0 + \Delta n} \tag{2.12}$$

where τ_{n0} and τ_{p0} are the respective capture time constants of electrons and holes and are defined as

$$\tau_{n0} := (N_T \sigma_n v_{th})^{-1} \qquad \text{and} \qquad \tau_{p0} := (N_T \sigma_p v_{th})^{-1} \tag{2.13}$$

with inverse dependencies on the thermal velocity, v_{th}, the defect concentration, N_T, and the capture cross-sections of the specific defect center in question, σ_n and σ_p [41].

This derivation is based on the assumptions of a non-degenerate semiconductor, a single defect level at a stable energy position, a negligible impurity concentration compared to the background doping, and a reemission time that is much larger than the relaxation time for captured carriers. The assumption of a single defect level can be reasonably extended to include multiple defect levels that dominate the carrier lifetime under different conditions, as long as there is negligible interaction between the levels. The last assumption effectively rules out the possibility of a defect acting as a trap center, which is not necessarily true for all conditions [41]. In [25], the SRH model is compared against experimental data and a more general model explicitly accounts for trapping effects, leading to the derivation of a "critical" recombination center density above which SRH theory diverges from experimental data. In the measurements presented in this thesis, the doping levels always remain below this critical doping level.

A more convenient expression for τ_{SRH} can be found by defining a symmetry factor k to be

$$k \equiv \frac{\sigma_n}{\sigma_p} = \frac{\tau_{p0}}{\tau_{n0}} \tag{2.14}$$

that depends only on the defect structure rather than on the absolute quantities of N_T and $\sigma_{n,p}$. Replacing τ_{p0} in Equation 2.12 with the symmetry factor, the SRH lifetime becomes

$$\tau_{SRH} = \tau_{n0}\left[\frac{p_0 + p_1 + \Delta n}{p_0 + n_0 + \Delta n} + k\frac{n_0 + n_1 + \Delta n}{p_0 + n_0 + \Delta n}\right] \tag{2.15}$$

Having introduced the symmetry factor k, the absolute defect parameters N_T and σ_n only appear in the electron capture time constant τ_{n0}, which is a common factor of both terms of τ_{SRH}. Consequently, τ_{n0} acts solely as a scaling factor for τ_{SRH} and does not affect the shape of SRH lifetime curves. Within the brackets of Equation 2.15, the relative defect parameters E_T and k only affect some of the terms, and therefore have a significant influence on the shape of various SRH lifetime curves (ie. temperature and injection dependence) [41].

2.3.1 Injection Dependence

The expression for τ_{SRH} given in Equation 2.15 can be simplified for low-level (LLI) and high-level injection (HLI) cases. For LLI ($\Delta n \ll n_0 + p_0$), the SRH lifetime becomes independent of the injection level,

$$\tau_{SRH}^{LLI} = \tau_{n0}\left[\left(\frac{p_0 + p_1}{p_0 + n_0}\right) + k\left(\frac{n_0 + n_1}{p_0 + n_0}\right)\right] \tag{2.16}$$

This can be further simplified by considering the cases of p-type and n-type semiconductors separately,

$$\tau_{SRH}^{LLI,p} = \tau_{n0}\left[\left(1 + \frac{p_1}{p_0}\right) + k\left(\frac{n_1}{p_0}\right)\right] \tag{2.17a}$$

$$\tau_{SRH}^{LLI,n} = \tau_{n0}\left[k\left(1 + \frac{n_1}{n_0}\right) + \left(\frac{p_1}{n_0}\right)\right] \tag{2.17b}$$

For a deep level trap, p_1 and n_1 are close in magnitude to the intrinsic carrier concentration. Consequently, the LLI SRH lifetimes from Equation 2.17 simplify to $\tau_{SRH}^{LLI,n} = \tau_{p0}$ and $\tau_{SRH}^{LLI,p} = \tau_{n0}$; since the low-level SRH lifetime is limited by the minority capture time constant, it is often called the minority carrier lifetime.

Similarly, for the high-level injection ($\Delta n \gg n_0 + p_0, n_1, p_1$), the SRH lifetime simplifies to become independent of both injection level and doping concentration,

$$\tau_{SRH}^{HLI} = \tau_{n0} + \tau_{p0} = \tau_{n0}(1 + k) \tag{2.18}$$

In the high-level case, the SRH lifetime is dependent only on the capture time constants and is therefore exclusively set by the trap density and and capture cross-sections; it is independent of the trap energy level [41].

2.3.2 Temperature Dependence

Considering p-type material, for example, and looking closely at Equation 2.17a, the temperature dependent terms are n_1, p_1, and τ_{n0}. The majority carrier concentration p_0 is also temperature dependent due to carrier freeze-out; this is critically important at cryogenic temperatures, but must also be considered across all temperatures for certain doping concentrations [45]. In the case of the lifetime measurements presented in this paper, however, freeze-out only needs to be considered for sufficiently low temperatures, due to the doping levels that are present. Assuming a trap center in the upper band gap half, the SRH lifetime reduces to

$$\tau_{SRH}^{LLI,p} = \tau_{n0}(T)\left[1 + k\frac{n_1(T)}{p_0}\right] \tag{2.19}$$

There are two contributions to the overall temperature dependence of the SRH lifetime: (1) τ_{n0}, which merely reflects the temperature dependencies of the capture cross-section σ_n and the thermal velocity v_{th}, and (2) the SRH density, n_1, which increases exponentially with increasing temperature (Equation 2.11). The temperature

dependence of v_{th} is given by

$$v_{th}(T) = v_{th}^{300\,\mathrm{K}}(T/300\,\mathrm{K})^{1/2} \tag{2.20}$$

whereas the temperature dependence of the n_1 stems from its dependence on the conduction band density of states N_C, expressed as

$$N_C(T) = N_C^{300\,\mathrm{K}}(T/300\,\mathrm{K})^{3/2} \tag{2.21}$$

The temperature dependence of σ_n depends entirely on the nature of the trap in question; in general the dependence follows a power law, but the variation between trap types is so great that this dependence may be increasing or decreasing with temperature.

At moderately low temperatures, the contribution of n_1 to the overall temperature dependence will be negligible, as its magnitude is far less than p_0. Therefore, the temperature dependence of the SRH lifetime is given directly by $\tau_{n0}(T)$, and is consequently proportional to the inverse product $(\sigma_n(T)v_{th}(T))^{-1}$. From this dependence, the superimposed dependence of the thermal velocity can be removed, resulting in the exact capture cross-section temperature dependence.

However, as the temperature increases and the Fermi level tends towards the middle of the band gap, the ratio n_1/p_0 will increase by orders of magnitude. When the Fermi level approaches the trap energy level, n_1/p_0 cannot be neglected, and eventually begins to dominate Equation 2.19, resulting in a steep increase in the SRH lifetime. From this, it is clear that the critical temperature for the onset of this steep lifetime increase is largely driven by the trap energy level. A shallow trap will manifest this increase at a much lower temperature than a deep level trap, due to its higher SRH density. Consequently, this onset temperature, along with a known background doping concentration, indicates the exact energy level of the trap. Moreover, the trap energy level can also be extracted from the slope of the lifetime for temperatures at which n_1/p_0 dominates the SRH lifetime equation.

2.4 Prior Work

To the author's knowledge, there exists little published temperature-dependent bulk lifetime data from commercial silicon-based technologies; one example can be found in [13], but it is limited to a temperature range of 300-400 K. Although the theoretical foundations of SRH theory are solid and long-established, and there have been many recent experimental and theoretical contributions by the photovoltaic community, these treatments have either been restricted to theory or have been confined to carefully arranged laboratory test cases designed to characterize the behavior of a particular impurity or defect. A significant amount of temperature and injection dependent lifetime data has been published from these tests, largely measured using photoconductive decay methods, such as the quasi-steady-state photoconductance (QSSPC) technique [50] and the microwave-detected photoconductance decay (UW-PCD) technique [5]. Because both of these techniques are non-contacting, they are well-suited for lifetime measurements on bare substrates and devices that are free of metal contacts. However, little has been done to comprehensively catalog the behavior of carrier lifetime across temperature within a commercially-available advanced silicon-based technology, particularly down towards the cryogenic temperature regime.

In the photovoltaic community, maximizing carrier lifetimes in the semiconductor bulk is critical to achieving high device efficiencies. To this end, extensive work has been done to characterize the recombination behavior of various impurities, such as copper [26], iron [7, 40], aluminum [44], and nickel [24], along with intrinsic defects such as a boron-related defect [27], oxygen-related defect [47], metastable defect in boron-doped Czochralski-grown silicon [39], and many of the radiation-induced defects described in Section 1.4.1 [8, 16]. One of the most common experimental techniques used to characterize the energy level and majority carrier capture cross-sections of recombination centers is deep level transient spectroscopy (DLTS), which is based on measurements of the transient capacitance of a diode [21]. However, although

DLTS allows the trap energy level to be precisely determined, it cannot be used to determine the minority carrier capture cross-section. For minority carrier applications, this weakness makes it difficult to predict the behavior of excess minority charge. Over the last decade, a new methodology that directly uses minority carrier lifetimes to characterize traps has been developed; many of the references above make use of these techniques, referring to them collectively as *lifetime spectroscopy* [46, 48, 49]. Lifetime spectroscopy addresses both the injection and temperature dependence of recombination lifetimes, and includes techniques that can use these dependencies in order to characterize many defect and material parameters, such as trap energy level E_T, trap density N_T, carrier capture cross-sections $\sigma_{n,p}(T)$, and background doping. When the temperature or injection dependence of lifetimes are treated separately, it is referred to as either *temperature dependent lifetime spectroscopy* (TDLS) or *injection dependent lifetime spectroscopy* (IDLS).

For the applications in this paper, however, the temperature dependence of the carrier lifetime within the bulk region of a commercial SiGe BiCMOS technology must be understood. Unlike the hands-on laboratory fabrication used to produce individual test devices with intentionally introduced impurities, the SiGe BiCMOS fabrication process involves side-by-side fabrication of advanced CMOS devices, SiGe HBTs, and a host of passive elements, along with multiple BEOL process steps. In order to incorporate accurate carrier lifetime models and develop predictive TCAD models for low-temperature and radiation-intense environments, lifetime characterization methodology must be applied to the actual technology that will be deployed. Since lifetime spectroscopy is in essence a set of analytical techniques for characterizing traps from minority carrier lifetime data, it is effectively decoupled from the actual measurement technique used to extract lifetime data. Consequently, since it is impossible to optically measure lifetimes in the bulk beneath a fully-processed HBT with metal contacts, all-electrical techniques can be employed instead, such as the

open circuit voltage decay measurement.

2.5 Description of Experimental Techniques

With the theoretical background just described, experimental measurements of SRH lifetime within a commercial SiGe BiCMOS process will be presented in the following sections. In this section, the experimental methods will be described, including overviews of the open-circuit voltage decay technique, the experimental setup, and the method for lifetime extraction.

2.5.1 Open Circuit Voltage Decay

The open circuit voltage decay (OCVD) technique is a well established method for measuring minority carrier lifetimes using semiconductor diodes [14]. Since the measurement is purely electrical, it is non-destructive and can be used to measure lifetimes in the bulk of the semiconductor. Other advantages include the ability to measure injection dependence, by varying the diode current bias, and temperature dependence, stemming from the ease of inserting diode test structures into cryogenic and high-temperature test fixtures.

The OCVD technique is based on the assumption of a step junction diode in which the wide of the lightly doped side is much larger than the minority carrier diffusion length. In this technique, the transient decay of the voltage across a diode is measured after instantaneously open-circuiting a diode that was passing forward current. In the simplest case, the transient decay exhibits a linear slope that is equal to $kT/(q\tau_B)$, where τ_B is the lifetime in the bulk (lightly-doped) region of the diode. The theoretical expression for the open circuit voltage for $V(t) \gg kT/q$ is [14]

$$V(t) = V_0 - (kT/q)\ \ln(\mathrm{erfc}(\sqrt{t/\tau_B})) \tag{2.22}$$

where $V(t)$ is the voltage across the diode as a function of time t, and V_0 is the instantaneous voltage after the diode is open-circuited. After differentiating $V(t)$ with

respect to time and simplifying for $t \gg \tau_B$, this equation reduces to

$$\frac{dV(t)}{dt} = -\frac{kT}{q\tau_B(1 - \tau_B/(2t))} \tag{2.23}$$

From these expressions, it is shown that after an initial nonlinear decay, the transient decay approaches a linear decay with a slope of $-kT/(q\tau_B)$.

As derived by Green [14], an expression for the transient voltage decay can be reached by using the minority carrier continuity equation for the base of the diode, given as

$$I_B = \frac{dQ_B}{dt} + \frac{Q_B}{\tau_B} \tag{2.24}$$

For the open-circuit measurement, $I_B = 0$, which simplifies the continuity equation to be $dQ_B/dt = -Q_B/\tau_B$. Using the chain rule,

$$\frac{dQ_B}{dt} = \frac{dQ_B}{dV} \cdot \frac{dV}{dt} \tag{2.25}$$

which after substituting into the continuity equation and solving for dV/dt results in

$$\frac{dV}{dt} = -\frac{Q_B}{\tau_B} \cdot \frac{dV}{dQ_B} \tag{2.26}$$

The excess minority carrier stored charge Q_B can be expressed as

$$Q_B = Aqn_{p0} = qA\frac{n_i^2}{N_B}\left(\exp\left(\frac{qV}{kT}\right) - 1\right) \tag{2.27}$$

Differentiating with respect to the diode voltage,

$$\frac{Q_B}{dV} = qA\frac{n_i^2}{N_B}\exp\left(\frac{qV}{kT}\right) \cdot \frac{q}{kT} \tag{2.28}$$

For $V \gg \dfrac{kT}{q}$, Equation 2.28 can be approximated as $\dfrac{dQ_B}{dV} \approx \dfrac{q}{kT}Q_B$. Substituting this expression into Equation 2.26, the resulting expression for dV/dt is given by

$$\frac{dV}{dt} = -\frac{kT}{q\tau_B}f \tag{2.29}$$

where f is given as

$$f = \frac{\dfrac{q}{kT}Q_B}{dQ_B/dV} \tag{2.30}$$

For the quasi-static approximation and $V \gg kT/q$, the factor f can be approximated as $f \approx 1$. Extracting dV/dt by fitting the slope of the measured linear voltage decay, the minority lifetime can be determined by rearranging Equation 2.29 as

$$\tau_B = -\frac{kT}{q}\left(\frac{dV}{dt}\right)^{-1} \quad , \; f \approx 1, \; V \gg \frac{kT}{q} \tag{2.31}$$

In practice, parasitics may exist in the measurement setup that will partially or entirely obscure the ideal linear voltage decay region. The first limiting parasitic is shunt capacitance, which is introduced either by measurement conditions or the inherent junction capacitance of the diode, and causes the voltage decay to pulled artificially high. The second parasitic to be considered is shunt resistance, which can be caused by unintentional leakage paths created during fabrication. Typical OCVD results are shown in Figure 2.2, including the ideal decay curve and both shunt capacitance and shunt resistance limited curves. These results were measured using a discrete diode (1N4001) with a standard lifetime of 6 µm. The ideal curve was measured without any modifications, while the capacitive and resistive curves were obtained by adding a shunt capacitance of 100 nF and shunt resistance of 74 kΩ, respectively.

Theoretically, parasitic resistance and capacitance can be addressed by accounting for both parasitics in the minority carrier continuity equation (Equation 2.24) and following the same derivation as given above. Including both a shunt resistance R_S and shunt capacitance C_S, the continuity equation becomes [14]

$$-\frac{V}{R_S} - C_S\frac{dV}{dt} = \frac{dQ_B}{dt} + \frac{Q_B}{\tau_B} \tag{2.32}$$

with corresponding solution

$$\frac{dV}{dt} = -\frac{kT}{q\tau_B} f \cdot \frac{\left(1 + \frac{V}{R_S}\cdot\frac{\tau_B}{Q_B}\right)}{\left(1 + \frac{fkT}{q}\cdot\frac{C_S}{Q_B}\right)} \tag{2.33}$$

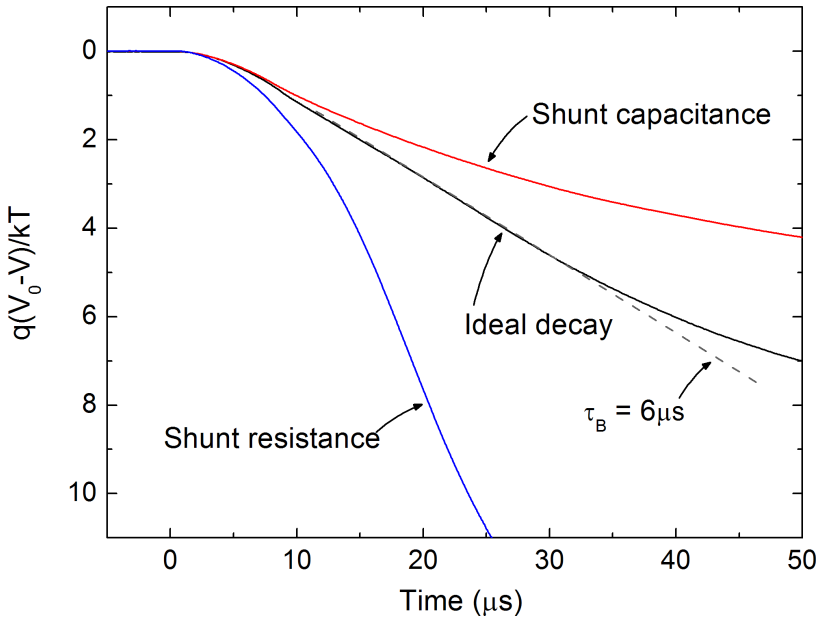


Figure 2.2: Open circuit voltage decay curves for a 1N4001 diode, showing an ideal decay, along with capacitance- and resistance-dominated curves.

From Equation 2.33, it is clear how a sufficiently large shunt capacitance will pull the decay curve up as it decreases the slope of the decay. Similarly, for a sufficiently small shunt resistance, the diode is partially shorted and the decay slope increases sharply. As seen in Figure 2.2, if either parasitic is dominating the decay, the linear decay region that is driven by the minority carrier lifetime is concealed altogether.

Conveniently, since the parasitics have opposing effects on the voltage decay, the detrimental effects of parasitic resistance or capacitance can potentially be removed. The smaller of the two parasitic terms in Equation 2.33 can be augmented by introducing additional shunt capacitance or resistance. In practice, parasitic leakage paths are not an issue for a mature commercial technology; instead, shunt capacitance typically poses the greatest challenge to OCVD measurements, due to the junction capacitance of the diode and the cable capacitances introduced by the measurement setup. If there is excessive capacitance in the measurement, then adding a shunt resistance will increase the numerator of Equation 2.33 and compensate for the capacitance; with proper tuning, it is possible to provide effective cancellation of the parasitics,

regaining an ideal decay curve and extracting the minority carrier lifetime.

2.5.2 Experimental Setup

The recombination lifetime measurements presented here include both injection-level and temperature variation. All measurements were conducted within IBM's first generation (5HP/5AM) and third generation (8HP) SiGe BiCMOS technologies, using custom designed diode test structures.

All DC measurements were conducted using an Agilent 4156 Semiconductor Parameter Analyzer. For the OCVD measurements, a Tektronix AFG3102 function generator was used to generate abrupt current pulses to the test structure by supplying a 1 kHz square wave through a 1N4148 fast-switching diode. A Tektronix TDS7054 oscilloscope was used to capture the transient voltage decay. Signal routing and variable shunt resistance compensation were provided by a custom PCB. Temperature dependent measurements of packaged test structures were carried out using a closed-cycle liquid-helium cryogenic test system capable of DC to 100 MHz operation from 4 K to 400 K. High temperature measurements were carried out using a Delta Design 9023 temperature chamber.

Fully automated measurements were made possible through the use of a Keithley

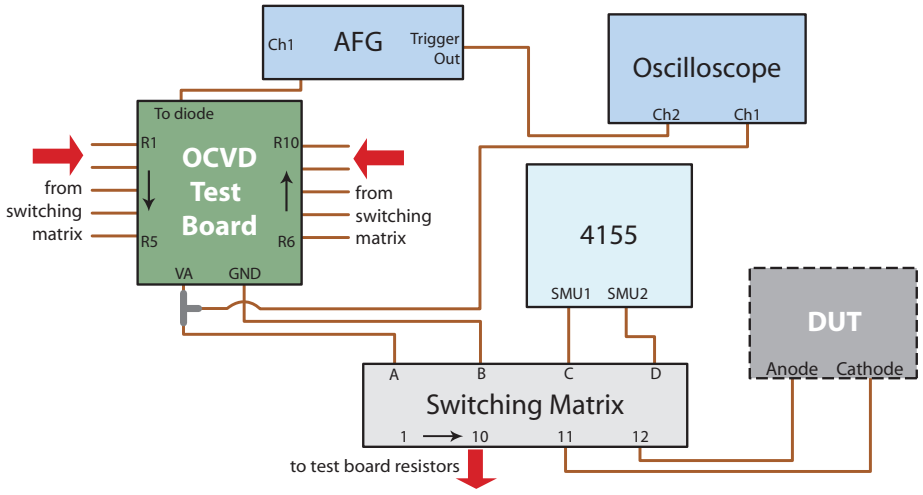


Figure 2.3: Diagram of the OCVD measurement setup.

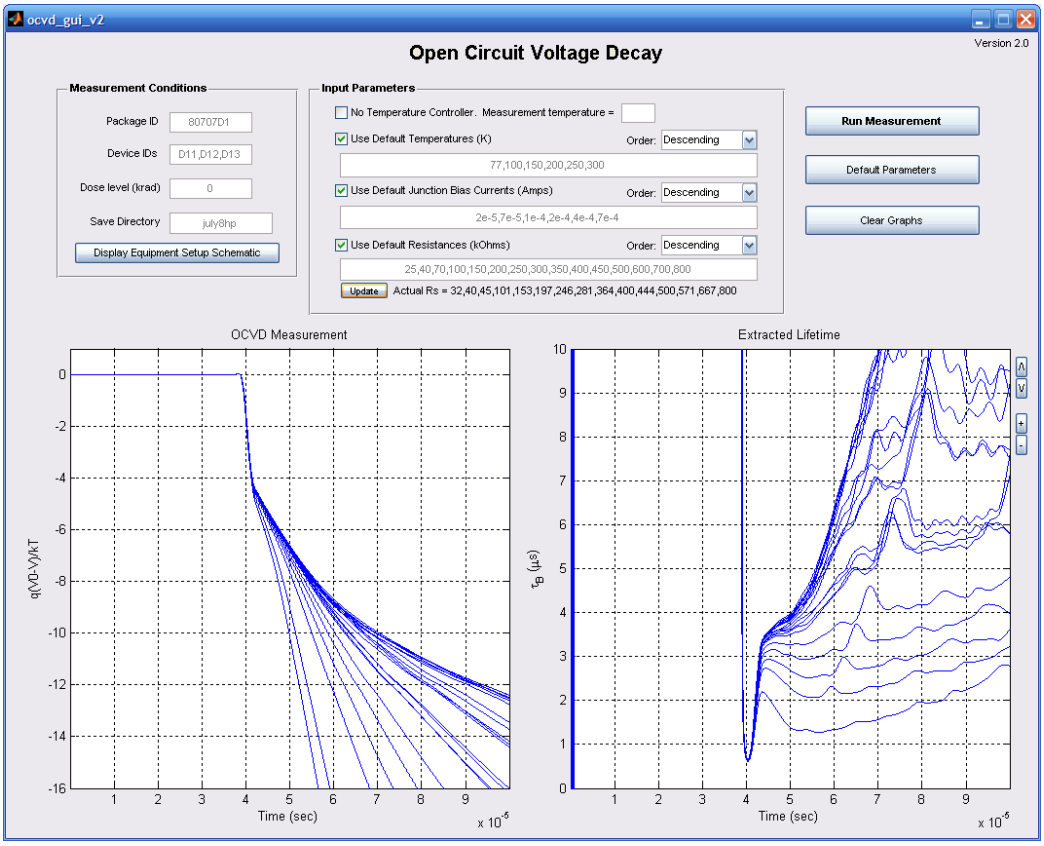


Figure 2.4: Automated OCVD measurement program.

7072 Switching Matrix and a MATLAB program that controls all equipment through the GPIB interface. A diagram of the experimental setup is given in Figure 2.3. This automated system makes it possible to efficiently and accurately measure carrier lifetimes across temperature and injection level. Since the specific shunt resistance needed to effectively cancel out parasitic capacitance is extremely sensitive and varies greatly with diode bias and temperature, at fixed temperature and bias, many measurements with varying R_S must be taken in order to reveal the lifetime. Thus, when varying both temperature and bias, the total number of measurements begins to be unwieldy without an efficient way to collect and process the data. The automated software used here can take as input sets of desired temperatures, current biases, and compensation resistances, stepping through all conditions without need for further user input.

A screen capture of the measurement program is given in Figure 2.4. The general measurement procedure consists of the following steps:

1. When initially reaching a specified temperature, measure the I-V characteristic of the diode test structure. Extract the applied voltages for the specified bias currents; this accounts for any series resistance and assures that the injection levels remain constant for all conditions.

2. Begin OCVD measurement, first without any applied shunt resistance. Perform a binary search with the input voltage level in order to reach the extracted diode voltage corresponding to the first desired current bias. When bias is reached, capture the transient waveform with the oscilloscope and transfer to computer.

3. Repeat step 2 for each specified shunt resistance. Calculate the combination of available shunt resistors that will provide the nearest equivalent resistance to each specified resistance.

4. Repeat step 2-3 for each specified current bias.

5. Repeat steps 2-4 for each specified temperature.

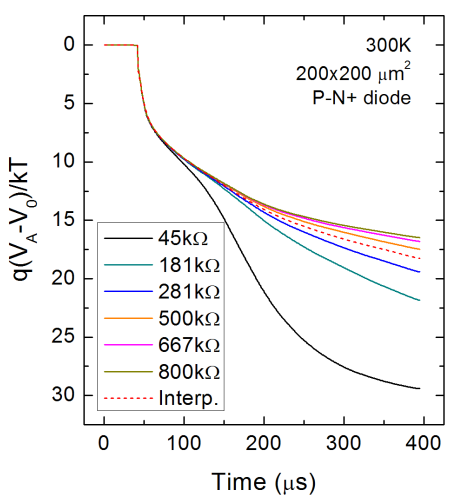
2.5.3 Lifetime Extraction

Having collected OCVD curves for all of the desired temperature and bias conditions, proper filtering and processing is critical to obtaining accurate carrier lifetime data. At many of the operating conditions, the voltage decay does not replicate the ideal result shown in Figure 2.2. This can result from excessive noise, very short lifetimes, non-ideal temperature effects, or imperfect compensation of parasitic capacitance. Furthermore, when the voltage decay is differentiated, inverted, and scaled to extract the lifetime, the effects of non-idealities are greatly magnified. Consequently, to remove these non-idealities, several tools can be used. To remove noise, the measured data can be filtered at various steps of the extraction. To account for imperfect compensation,

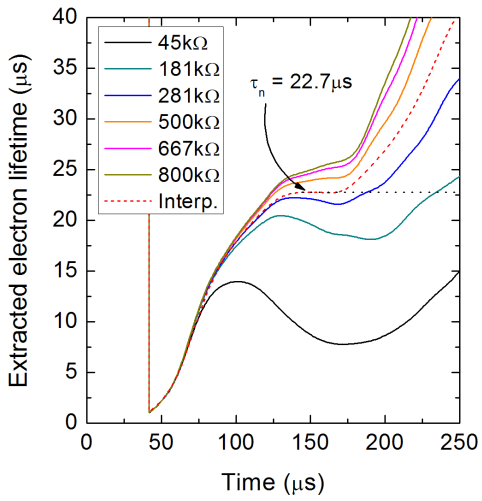
the filtered decay curves with varied R_S can be interpolated to predict the ideal linear decay and extract a precise and accurate lifetime value.

When discussion lifetime extraction, we refer to the data as a set of OCVD curves for varied R_S and a fixed temperature and bias. The best approach found for filtering is to first pass the data through a Butterworth low-pass filter (10-100 kHz) in order to remove the worst of the high-frequency noise, then to pass the data through a Gaussian filter. When considering various filters, we found that low-pass filters with too low of a cutoff frequency and averaging functions with too wide of a span tended to distort the magnitude of the carrier lifetime.

After the initial filtering, the data can optionally be interpolated. Subsequently, the data is differentiated, then inverted to result in a set of lifetime extraction curves. Low-pass and Gaussian filters are again applied after each of these steps. Finally, since these curves represent the inverse derivative of the voltage decay, a period of time with a constant lifetime indicates a linear voltage decay and corresponds to the carrier lifetime. An example set of measured OCVD and lifetime extraction curves is shown in Figure 2.5, with the dotted line indicating the ideal interpolated result. The validity of interpolation was verified by comparing the lifetime extraction curve from a measured OCVD curve at a particular R_S with a lifetime extraction curve generated from an OCVD curve that was interpolated from the rest of varied-R_S OCVD curves. A set of interpolated curves generated from the measurements in Figure 2.5 is shown in Figure 2.6.

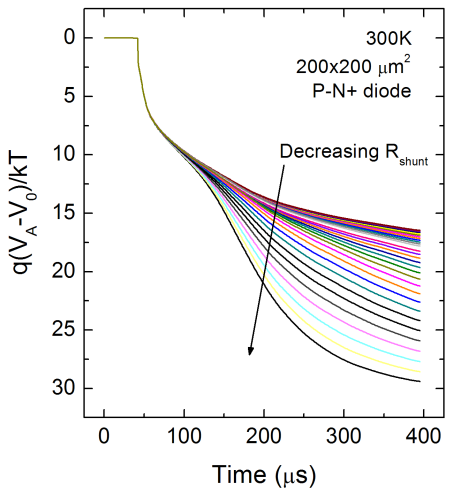


(a) Example set of OCVD curves.

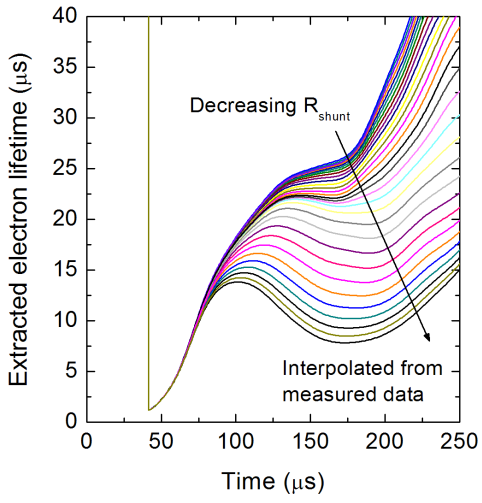


(b) Matching set of lifetime extraction curves.

Figure 2.5: Example of the lifetime extraction method used in this work. In (a), a set of OCVD curves for varying shunt resistance is shown. The corresponding lifetime extraction curves are shown in (b), with the dotted curves indicating the interpolated ideal OCVD characteristic, exhibiting the linear decay indicative of the minority carrier lifetime.



(a) Example set of interpolated OCVD curves.



(b) Matching set of lifetime extraction curves.

Figure 2.6: Example of data interpolation used to extract carrier lifetime. In (a), a set of interpolated OCVD curves for varying shunt resistance is shown. The corresponding lifetime extraction curves are shown in (b).

2.6 Experimental Results

Following the above procedure, the minority electron lifetime in the substrate of IBM's 8HP SiGe BiCMOS technology was measured using custom diode structures. In order to measure the electron lifetime, p-n+ diodes were constructed from the substrate-subcollector junction, similar to Figure 1.2. Multiple samples of both 100x100 μm^2 and 200x200 μm^2 sizes were measured. Similarly, the minority hole lifetime was measured using p+n- diodes built from the p+ source/drain diffusion within the lightly-doped n-well region. Layouts for these test structures are given in Figure 2.7.

In Figure 2.8, the injection dependence of the minority electron lifetime is shown across a range of temperatures. The corresponding temperature dependence is given in Figure 2.9, in which the bias current is fixed in order to decouple the lifetime injection dependence. Measurements from 50 K to 350 K were conducted in the cryogenic test system, whereas measurements from 325 K to 500 K were conducted using the temperature chamber. The OCVD curves were found to depart entirely from the expected behavior at exceedingly low or high temperatures; thus, the electron lifetimes presented here are limited to a temperature range of 90 to 425 K.

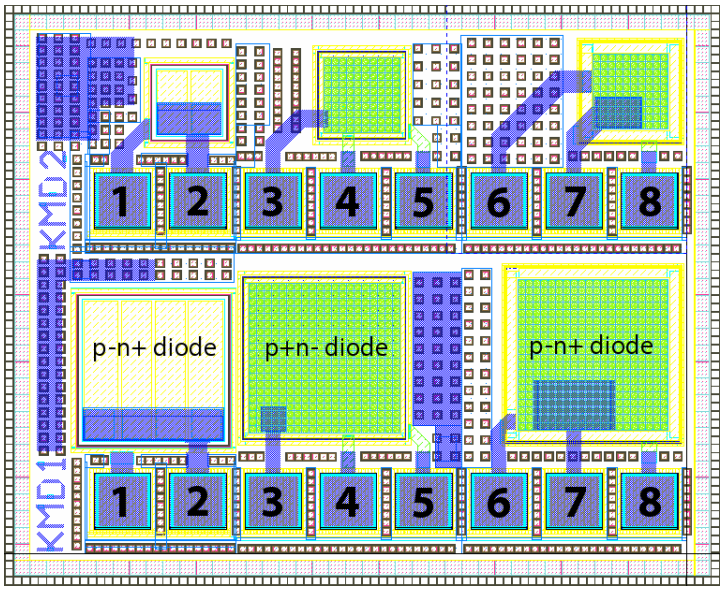


Figure 2.7: Layout of diode test structures used for lifetime measurement.

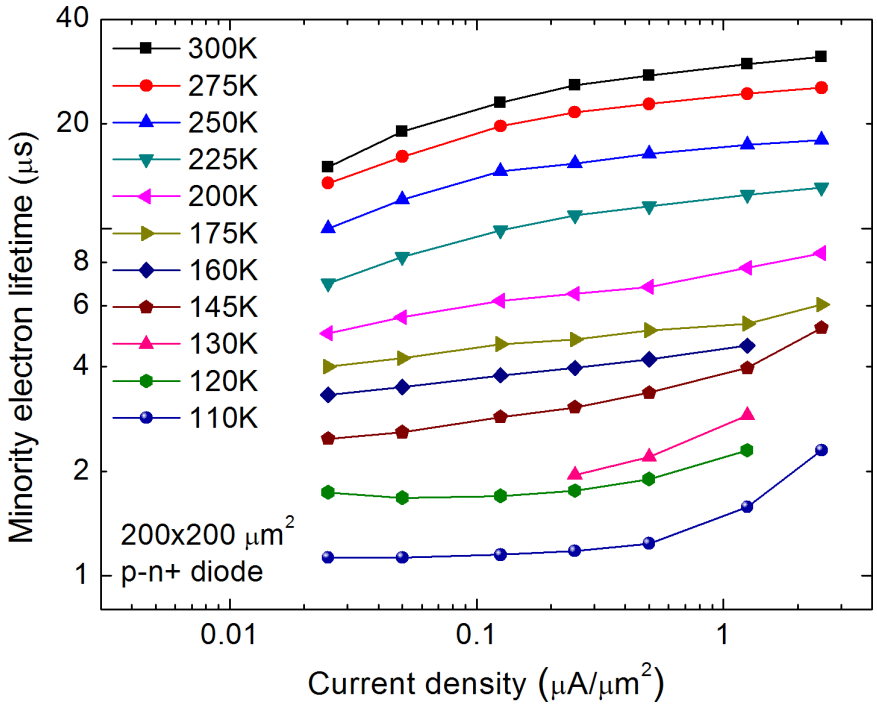


Figure 2.8: Minority electron lifetime across temperature as a function of diode current bias.

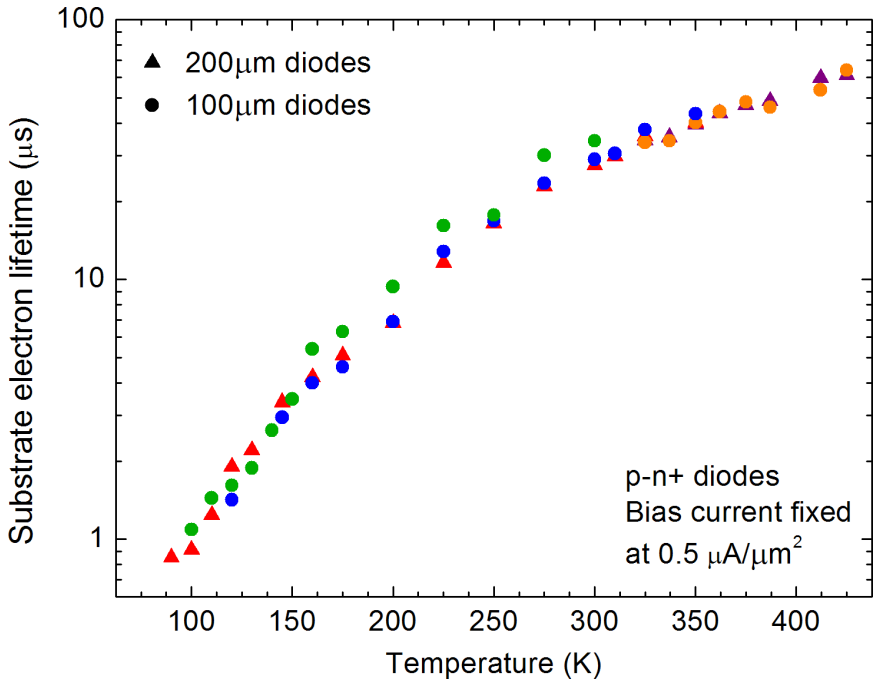


Figure 2.9: Minority electron lifetime across temperature as extracted from multiple test samples. Circles indicate 100x100 μm^2 diodes, whereas triangles indicate 200x200 μm^2 diodes.

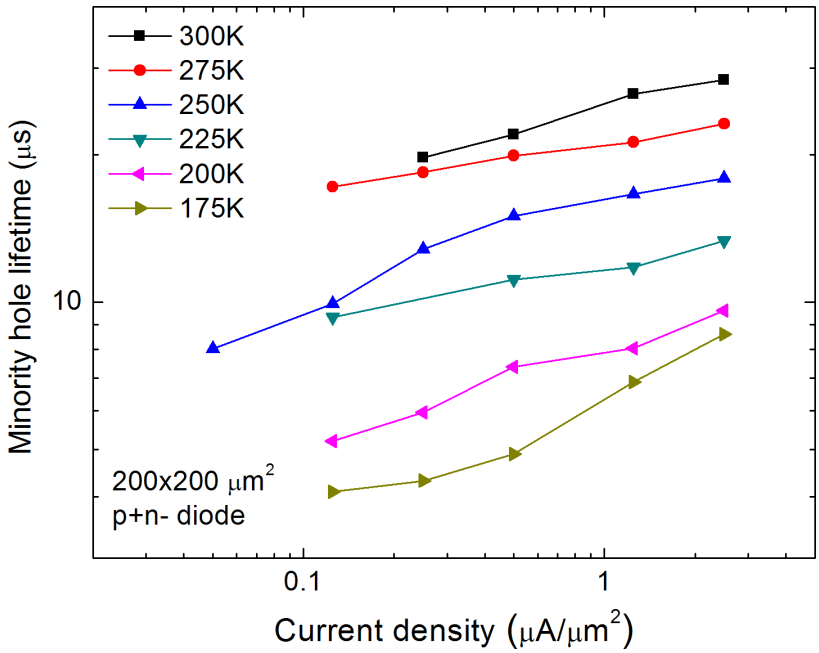


Figure 2.10: Minority hole lifetime across temperature as a function of diode current bias.

In Figure 2.10, the injection dependence of the minority hole lifetime is shown across a range of temperatures. The corresponding temperature dependence is given in Figure 2.11, in which the bias current is fixed in order to decouple the lifetime injection dependence.

The fact that both the electron and hole lifetimes increase with increasing temperature indicates that their respective capture cross-sections have an inverse dependence on temperature. The lack of a rapid increase in lifetime up to a temperature of 425 K indicates that the dominant trap energy level fairly deep or close to the middle of the band gap. These issues are further discussed in the following section.

Proton irradiation experiments were conducted at the Crocker Nuclear Laboratory at UC Davis in order to assess the effects of displacement and ionization damage of the minority carrier lifetimes. The diode test structures were subjected at room temperature to 63 MeV proton irradiation up to a total accumulated dose of 1 Mrad (Si).

Figure 2.12 shows the radiation response of the I-V characteristic of the p-n+

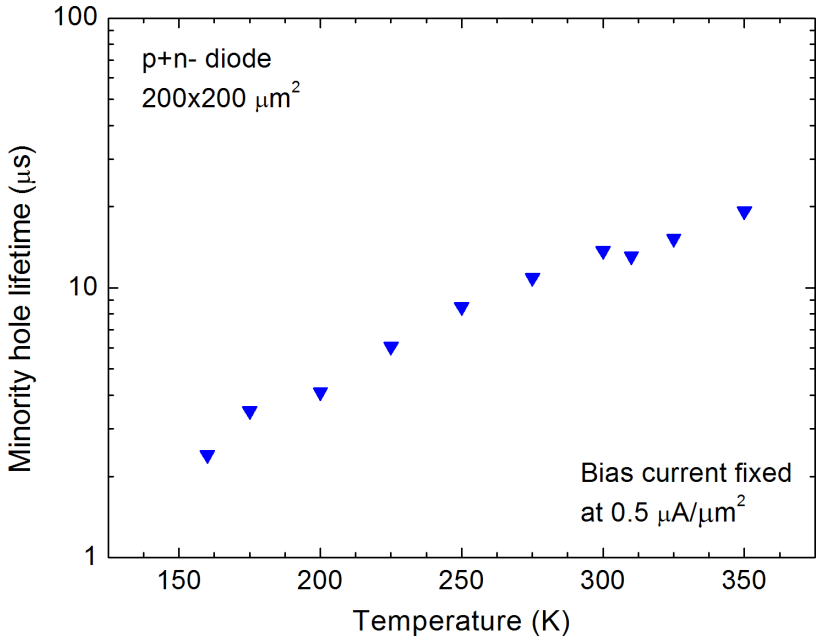


Figure 2.11: Minority hole lifetime across temperature.

diode used for electron lifetime measurements. Prior to irradiation, the diode I-V curve exhibits a nearly ideal slope, as expected for a properly fabricated diode. With increasing proton dose, the slope of I-V curve increasingly departs from the ideal slope of 60 mV/decade due to increased recombination. The resulting minority electron lifetime injection dependence across proton dose is shown in Figure 2.13. Temperature dependent measurements were also carried out, with the results given in Figure 2.14.

Likewise, Figure 2.15 shows the radiation response of the I-V characteristic of the p+n- diode used for hole lifetime measurements. Prior to irradiation, the diode I-V curve also exhibits a nearly ideal slope. With increasing proton dose, the slope of I-V curve increasingly departs from the ideal slope of 60 mV/decade due to increased recombination. The measured minority hole lifetime injection dependence across proton dose is shown in Figure 2.16. Temperature dependent measurements were also carried out, with the results given in Figure 2.17. As expected intuitively and from the I-V characteristics, both hole and electron lifetimes decrease substantially due to increased recombination induced by displacement and ionization damage.

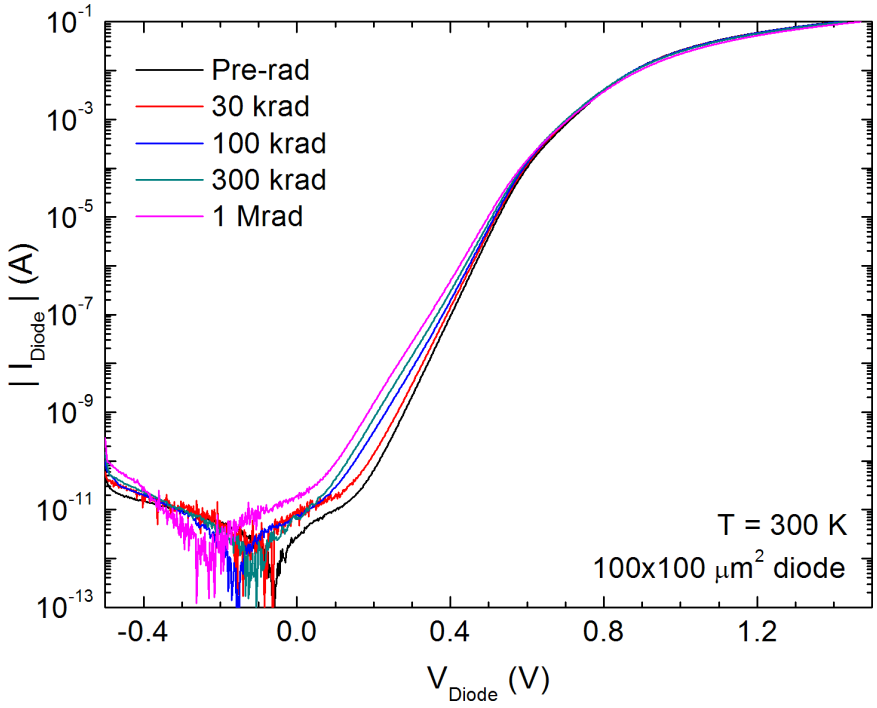


Figure 2.12: Radiation response of p-n+ diode I-V characteristic.

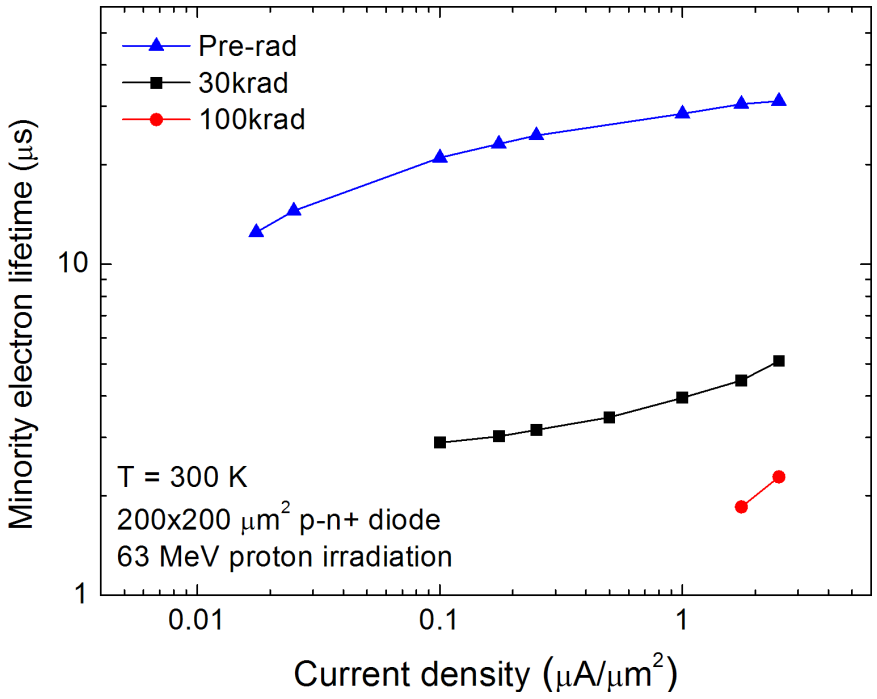


Figure 2.13: Radiation response of minority electron lifetime as a function of bias.

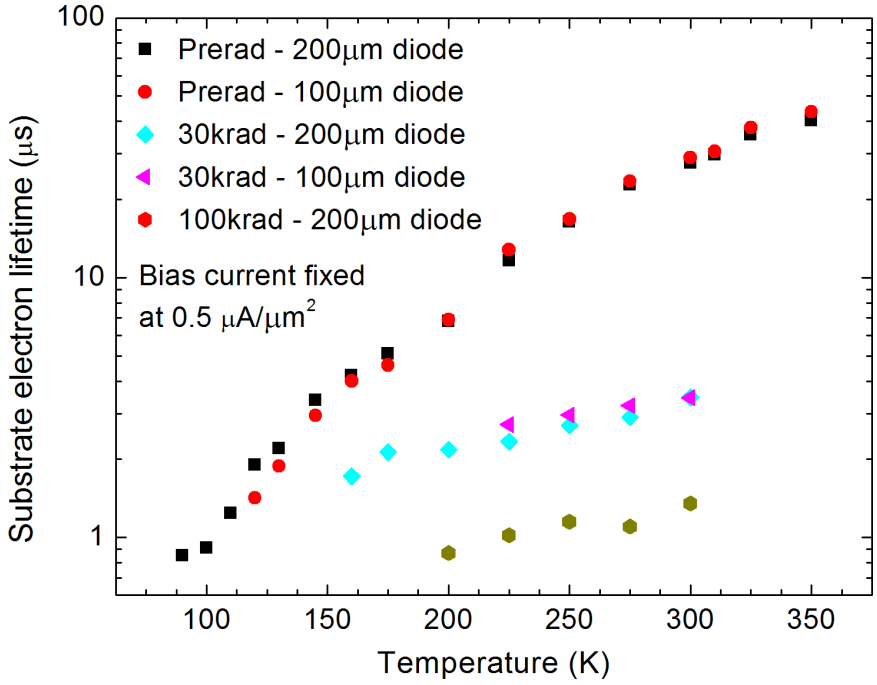


Figure 2.14: Radiation response of minority electron lifetime across temperature.

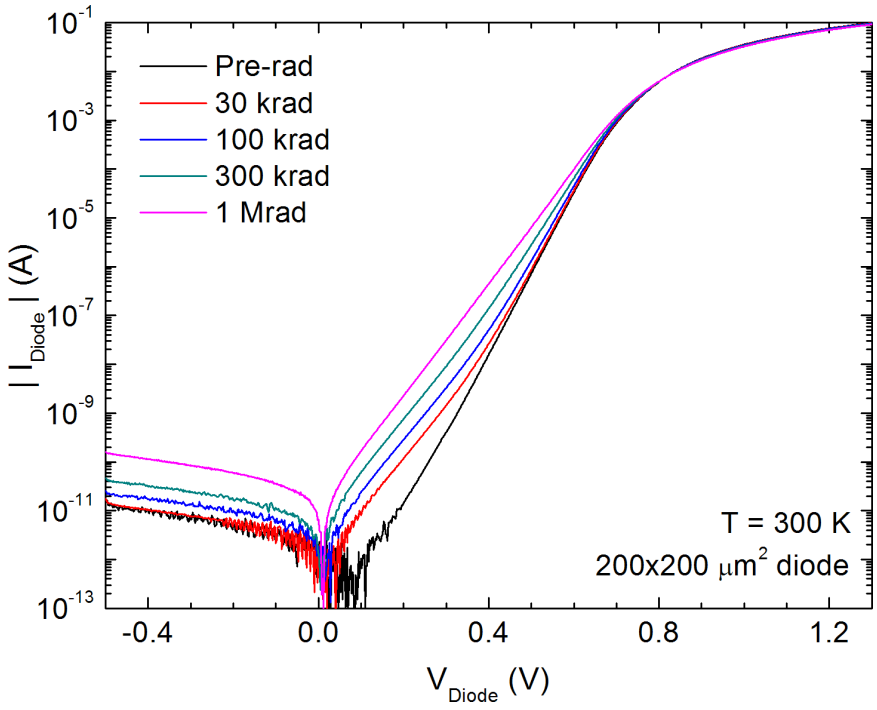


Figure 2.15: Radiation response of p+n- diode I-V characteristic.

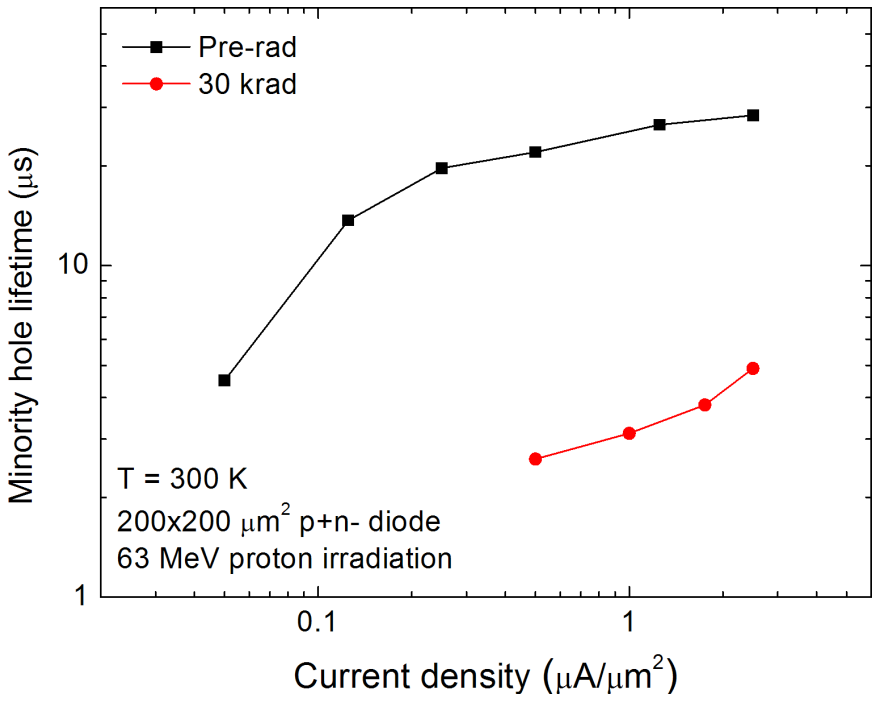


Figure 2.16: Radiation response of minority hole lifetime as a function of bias.

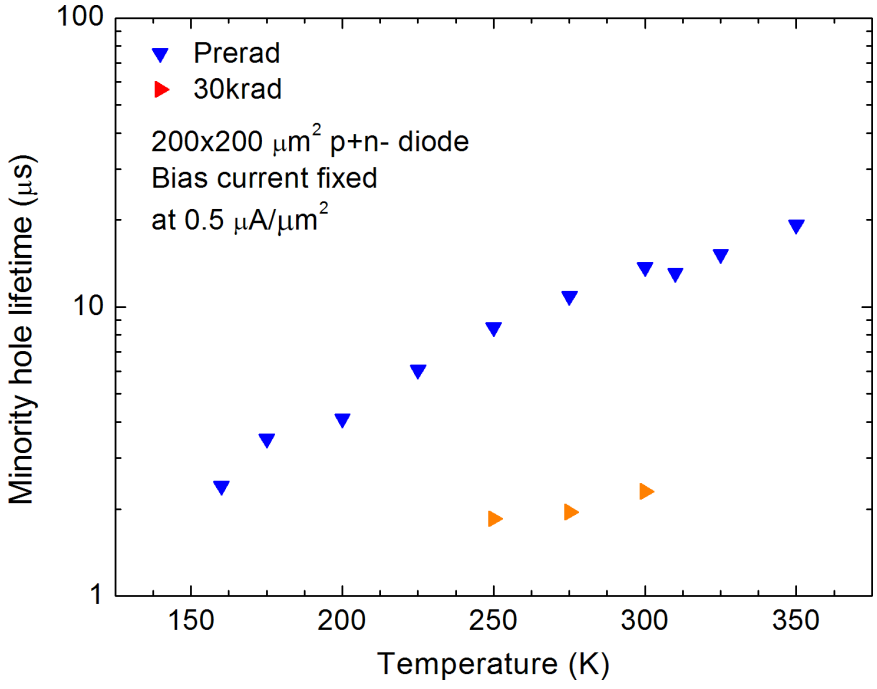


Figure 2.17: Radiation response of minority hole lifetime across temperature.

2.7 Modeling of Data

The temperature and injection dependent data for minority electron SRH lifetimes that is presented in the previous section can be used to develop calibrated theoretical models that can be inserted into commercial TCAD software. Considering this data in light of the SRH recombination theory presented in Section 2.3 and the analytical techniques of lifetime spectroscopy, it is possible to characterize the dominant trap levels within this particular technology.

The trap parameters for a selection of common defects and impurities are given in Table 2.1, including an arbitrary trap density chosen in order to generate the theoretical temperature dependence for each defect (Figure 2.18). From Figure 2.18, it is clear that the temperature dependencies of the defects and impurities have widely varying shapes. As described in Section 2.3, at a sufficiently high temperature the slope of the lifetime sharply increases due to the rising SRH density n_1 or p_1. The specific onset temperature of this effect is determined by a combination of the trap energy level and the background doping concentration, whereas the slope of this increase is chiefly determined by the trap energy level. As in the case of Ti (a fairly shallow defect), the onset temperature is near room temperature, whereas for deep traps the steep

Table 2.1: List of some common impurities and defects, along with published values for their defect parameters. The specified trap densities were arbitrarily chosen and are used to generate the theoretical curves shown in Figure 2.18.

Trap	Energy [eV]	$\sigma_p(T)$ [cm^2]	$\sigma_n(T)$ [cm^2]	N_T [cm^{-3}]
E1 [16]	E_C - 0.164	$6.4\times10^{-7}\times\exp(\frac{-T}{150})/v_{th}$	$8\times10^{-8}\times T^{0.7}/v_{th}$	10^{11}
E4 [16]	E_C - 0.421	$5.4\times10^{-9}\times T^{0.4}/v_{th}$	$2\times10^{-6}\times T^{-0.3}/v_{th}$	10^{10}
Ni [41]	E_C - 0.4	$8\times10^{-17}\times(T/300)^{-2.4}$	$5.6\times10^{-17}\times(T/300)^{-2.4}$	10^{14}
Mo [41]	E_V + 0.317	$6\times10^{-16}\times(T/300)^{-1.6}$	$7.8\times10^{-15}\times(T/300)^{-1.6}$	10^{11}
Ti [41]	E_V + 0.289	$1.9\times10^{-16}\times(T/300)^{-1.6}$	$2.3\times10^{-15}\times(T/300)^{-1.6}$	10^{11}
Cz [41]	E_C - 0.411	$1\times10^{-15}\times(T/300)^{-2}$	$9.5\times10^{-15}\times(T/300)^{-2}$	10^{11}
Cu [41]	E_C - 0.58	$1\times10^{-16}\times(T/300)^{-2}$	$1\times10^{-15}\times(T/300)^{-2}$	10^{11}

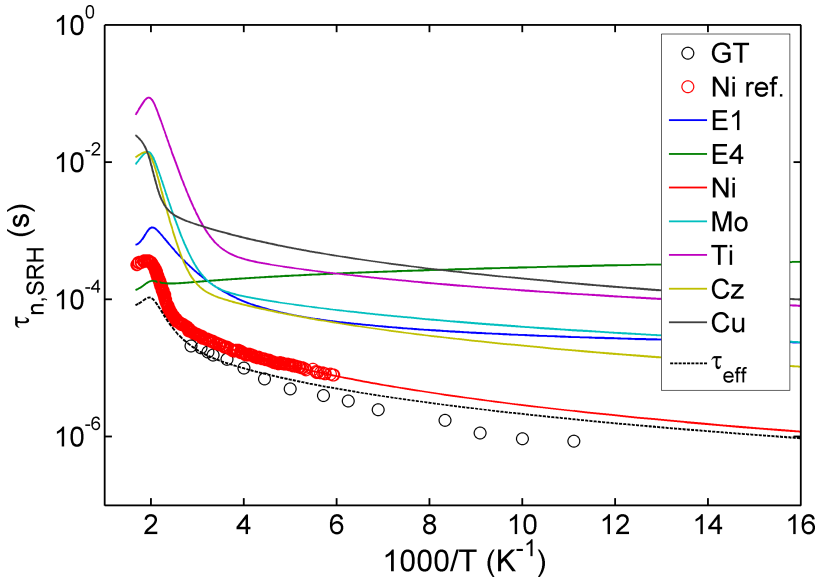


Figure 2.18: SRH lifetime model calculations for various defects and impurities. The red circles indicate lifetimes from a Ni-contaminated sample, with an excellent fit (red line) by the theoretical SRH model [41]. The data presented in this work is included for comparison.

increase may not appear until 450 K or higher. From the lower temperature regime, the shape of the lifetime curve is entirely driven by the temperature dependence of the capture cross-sections. As a result, with data covering a sufficient temperature range, $\sigma_n(T)$ and $\sigma_p(T)$ can be easily extracted.

First, modeling of the minority electron lifetime will be considered. Because the minority electron lifetimes reported in this paper did not exhibit a steep high temperature increase, the dominant trap in the substrate must be a deep energy level trap with an onset temperature above 425 K. The temperature and injection dependent lifetime data were simultaneously fit by using the SRH model and varying trap energy, trap density, $\sigma_n(T)$, and $\sigma_p(T)$. The resulting parameters are listed

Table 2.2: Trap parameters extracted from minority hole and electron lifetime data.

Trap	Energy [eV]	$\sigma_p(T)$ [cm^2]	$\sigma_n(T)$ [cm^2]	N_T [cm^{-3}]
GT_n	E_C - 0.5	$1\times10^{-15}\times(T/300)^{-3.49}$	$3.5\times10^{-15}\times(T/300)^{-3.49}$	4.5×10^{12}
GT_p	E_V + 0.32	$3.1\times10^{-15}\times(T/300)^{-3.05}$	$3.1\times10^{-15}\times(T/300)^{-3.05}$	4.3×10^{12}

in Table 2.2, where the trap is labeled as GT_n. With this data, it is important to note that only the combined magnitude of the $N_T \cdot \sigma_n$ product can be extracted, since the magnitude of the lifetime is dependent on the product of these parameters. Additional information is needed in order to decouple these two parameters. However, since their relative magnitudes do not affect the shape of the lifetime injection and temperature dependence, the trap densities for GT_n and subsequent defect models have been arbitrarily chosen. Figure 2.19 demonstrates the accuracy of the fit across the entire temperature range. A trap energy level of 0.5 eV below E_C was used for this calculation; it was found that when the trap energy level was within 0.4 eV of the conduction or valence bands, the high temperature fit began to diverge substantially from the data due to the onset of the steep lifetime increase.

The quality of the injection dependent lifetime fit using the GT_n model is shown in Figure 2.20 for both 225 K and 300 K. The injection level is defined as the ratio of the injected minority carrier density to the majority carrier concentration (neglecting freeze-out, $\Delta n/N_A$). In order to estimate the injected carrier density, the diode voltage was measured immediately after open-circuiting the diode in order to remove excess voltage drop due to series resistance. Using this actual applied voltage and known doping concentration, the injected carrier concentration was determined from the following boundary condition: $\Delta n_p = n_i^2/N_A(\exp(qV_A/kT) - 1)$ [37]. For a current density of 1 $\mu A/\mu m^2$, the actual diode voltage is 0.665 V, resulting in an injected carrier density of approximately $1.8 \times 10^{16}\,cm^{-3}$, which translates to an injection level of 20.13.

Similar to the electron lifetime, the dominant trap indicated by the minority hole lifetime in the n-well is a mid to deep energy level trap. The extracted trap parameters are given in Table 2.2, where the trap is labeled as GT_p. In Figure 2.21, the resulting temperature dependent model is plotted against the measured lifetimes. A trap energy level of 0.32 eV above E_C was used for this calculation; it was found that when the

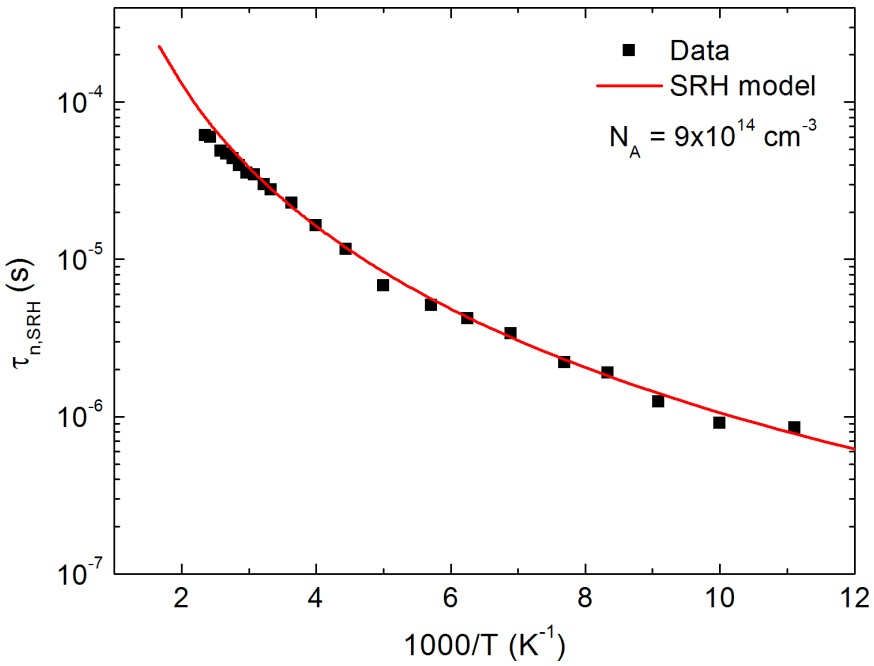


Figure 2.19: SRH model fit of minority electron lifetime vs. temperature.

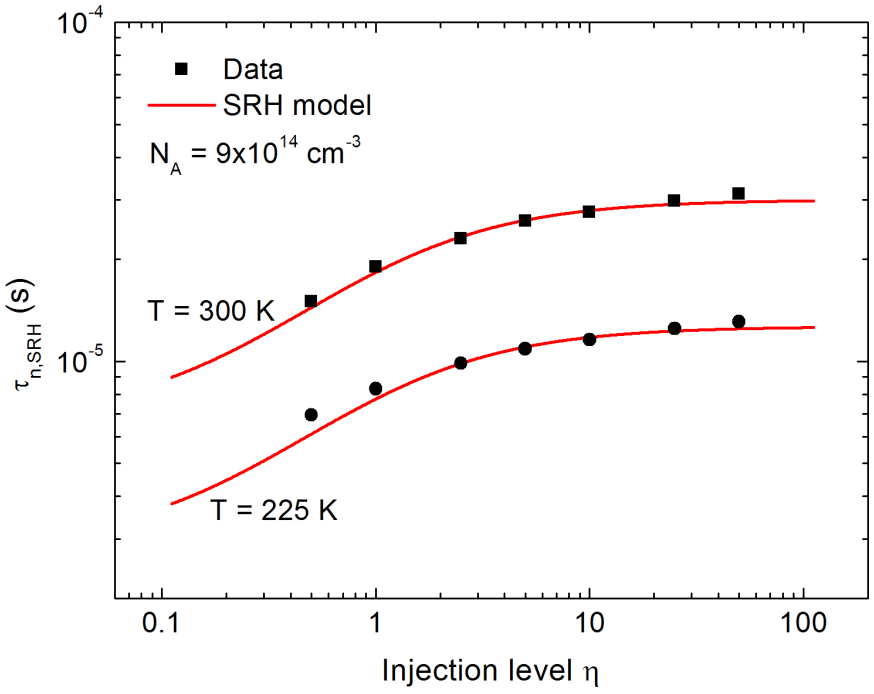


Figure 2.20: SRH model fit of minority electron lifetime vs. injection level. Injection level η is specified as the ratio of the injected minority carrier density to the majority carrier concentration.

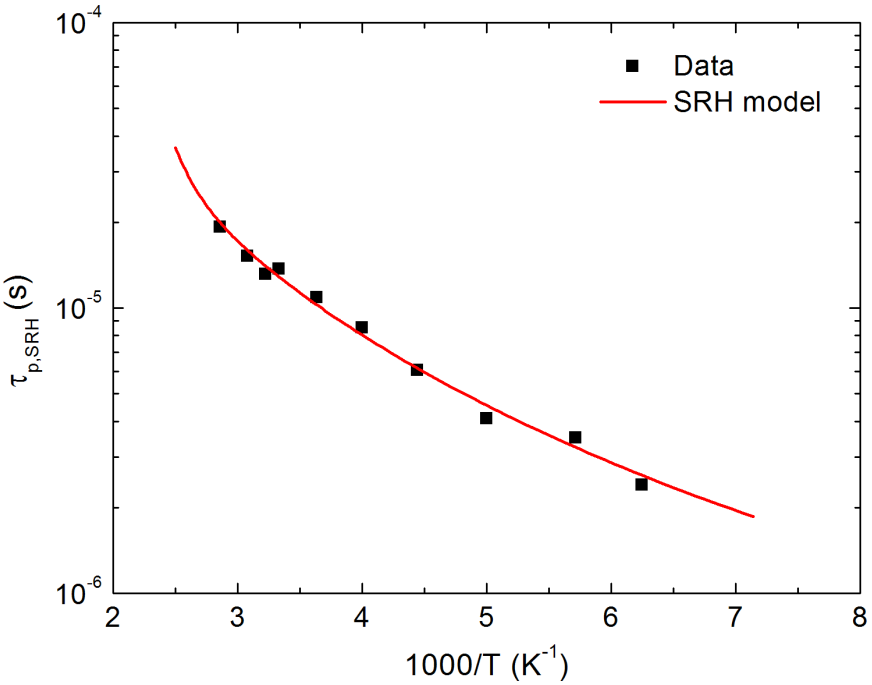


Figure 2.21: SRH model fit of minority hole lifetime vs. temperature.

trap energy level was within 0.3 eV of the conduction or valence bands, the high temperature fit began to diverge substantially from the data due to the onset of the steep lifetime increase.

Finally, the irradiated electron lifetime data can be modeled by introducing additional trap energy levels. As discussed in [16], the primary energy levels produced by electron irradiation in p-type silicon are the E1 and E4 defects. The E1 defect is a vacancy-oxygen complex, whereas the E4 defect is a divacancy [16]. By maintaining the same GT_n trap density and increasing the trap densities of the E1 and E4 levels, it was possible to closely fit the measured temperature dependence of the 30 krad sample, shown in Figure 2.22. The resulting trap densities used in this model are $N_{T,E1} = 1 \times 10^{12}\,\mathrm{cm}^{-3}$ and $N_{T,E4} = 4.2 \times 10^{12}\,\mathrm{cm}^{-3}$.

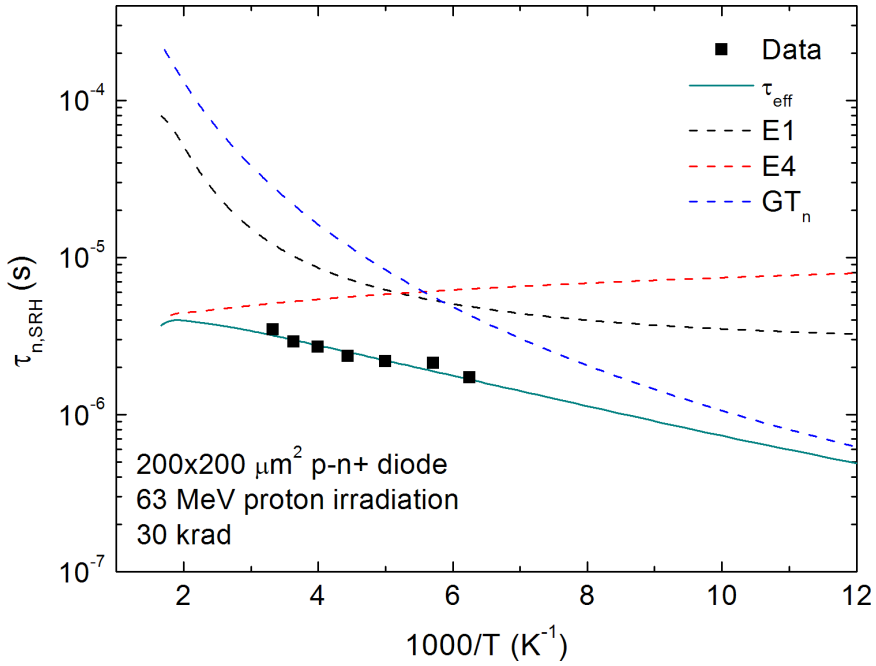


Figure 2.22: SRH model fit of minority electron lifetime vs. temperature after 30 krad proton irradiation. The effective lifetime is determined by the inverse sum of the reciprocal lifetimes of all contributing traps.

2.8 Summary

In this chapter, fundamental recombination mechanisms were reviewed, and a careful analysis of SRH recombination theory was presented, leading to a theoretical understanding of the temperature and injection dependence of SRH lifetimes. Prior work addressing the temperature dependence of SRH lifetime was reviewed, revealing both sophisticated analytical methods and a significant amount of temperature and injection dependent lifetime data to which those methods can be applied. However, the dearth of experimental data that addresses the temperature dependence of lifetimes within a advanced silicon-based technology renders it impossible to accurately predict the behavior of these lifetimes, since the behavior of carrier lifetimes is wildly dependent on the specific trap centers within a particular technology. A sophisticated and advanced technology will likely have a significantly different composition of traps than a carefully arranged laboratory test sample that has undergone minimal processing.

With this in mind, new experimental data characterizing the temperature and injection dependence of carrier lifetimes within a commercial SiGe BiCMOS technology was presented. In order to further develop accurate lifetime models that can be applied to extreme environments with high levels of radiation, proton irradiation tests were conducted, followed by temperature dependent lifetime measurements. The collection of experimental data was analyzed and lifetime spectroscopy techniques were applied in order to extract the dominant trap parameters, leading to minority electron and hole lifetime models that are compatible with commercial TCAD software.

CHAPTER III

LOW-TEMPERATURE RESISTIVITY IN SILICON

In this chapter, an overview is given of the approaches taken to model the temperature dependence of resistivity and mobility down to the deep cryogenic regime. The predominant models for mobility and incomplete dopant ionization are addressed, then applied to experimental resistivity data taken from both p-type and n-type samples of various doping concentrations. By calibrating these models, improved fits of the experiments are achieved across doping concentration and temperature. In addition, the radiation response of substrate resistivity is evaluated experimentally.

3.1 Introduction

Mobility is the critical parameter in determining the behavior of electron and hole transport due to drift fields, since for low to moderate electric fields it is defined as the constant of proportionality between the carrier drift velocity and the electric field acting on that carrier. As a result, the carrier mobility plays a key role in determining the performance of most semiconductor devices. For example, Equation 1.6 shows that the collector current of the SiGe HBT is directly dependent on the base electron mobility, since the base diffusion constant D_{nb} is related to the mobility through the Einstein relation, $D_{nb} = \frac{kT}{q}\mu_{nb}$.

Whereas mobility measures the degree to which carriers are able to freely move through the crystal lattice, its converse indicates the degree to which carriers undergo momentum-decreasing collisions within the crystal lattice. The frequency of these *scattering mechanisms* varies inversely with the carrier mobility, following the general relation $\mu = q\langle\tau\rangle/m*$, where $\langle\tau\rangle$ is the mean free time between collisions and $m*$ is the carrier effective mass [37]. In a device-quality nondegenerately doped semiconductor,

the scattering mechanisms are can be categorized as (1) lattice scattering involving collisions with thermally agitated lattice atoms, (2) majority impurity scattering, (3) minority impurity scattering, and (4) electron-hole scattering [17].

Clearly, the mobility is heavily dependent on parameters such as the doping concentration, since scattering increases with a greater number of impurities, and the temperature, since a higher lattice temperature leads to systematically increased thermal agitation of lattice atoms. Similarly, the mobility is tightly coupled to other phenomena, such as incomplete ionization of majority impurities, since ionized dopants will necessarily interact differently with charge carriers than inactive dopants. High carrier density effects must also be addressed, due to the fact that at high carrier concentrations, carriers tend to screen impurities from other carriers [18].

In order to calibrate a mobility model that is accurate at temperatures down to the deep cryogenic regime, experimental data will be presented in this chapter. However, direct measurements of mobility are difficult to achieve, similar in some respects to the challenges posed to carrier lifetime measurements. Ideally, it would be possible to calibrate a mobility model using measurements from simple test structures fabricated within the semiconductor technology to which the model will be applied. The simplest and most robust measurement that can be carried out with a direct relevance to mobility is the kelvin measurement. Simple test structures can be used, such as a four-point-probe structure to measure resistivity or a ring-dot structure to measure sheet resistance. However, the problem inherent to resistance measurements is that the behavior of the mobility is necessarily married to that of the majority carrier concentration, since the resistivity of the material follows the relation

$$\rho = \frac{1}{q\mu_p N_A}, \qquad \text{for p-type material} \tag{3.1}$$

Particularly at low temperatures, the majority carrier concentration can be significantly reduced due to incomplete ionization. However, despite the challenges posed by the mingling of both ionization and mobility effects, with a wide range of resistivity

measurements for p-type and n-type silicon of varied doping concentrations it is possible to gain insight into the low temperature behavior of both incomplete ionization and carrier mobility. Calibrated models for both phenomena can then be developed and applied towards accurate predictive TCAD simulations of SiGe devices and circuits. Since incomplete ionization also plays a significant role in device operation, calibrating an ionization model in addition to a mobility model will provide further benefits to the overall field of SiGe device simulation.

3.2 Resistivity Modeling

In this section, the current state-of-the-art simulation models for mobility and incomplete ionization will be reviewed, highlighting their key features and analyzing their treatment of temperature effects. Specifically, the two models that will be described are the physics-based Philips unified mobility model described by Klaassen [17, 18, 19] and the model for density of states and incomplete ionization proposed by Altermatt [2, 3].

3.2.1 Carrier Mobility

The Philips mobility model is a physics-based analytical model and is referred to as a unified mobility model because it unifies the descriptions of the majority and minority carrier mobilities. Besides lattice, donor, and acceptor scattering, this model also incorporates the effects of impurity screening by charge carriers, electron-hole scattering, clustering of impurities [18], and a full temperature dependence for both majority and minority carrier mobility [19]. Moreover, since the model gives the carrier mobility as an analytical function of the donor, acceptor, electron, and hole concentrations, it is a natural fit for implementation within a device simulator.

For the parameter descriptions and values used in the mobility model equations that are presented in this section, refer to Klaassen [18, 19]. At room temperature, the majority electron and hole mobilities are well characterized by experimental data

across a wide range of impurity concentrations, following the expression

$$\mu = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + (N/N_{ref,1})^{\alpha_1}} - \frac{\mu_1}{1 + (N_{ref,1}/N)^{\alpha_2}} \tag{3.2}$$

At low impurity concentrations, the dominant scattering mechanism is lattice scattering; thus, the low-concentration limit of Equation 3.2 describes the room temperature carrier mobility due to lattice scattering,

$$\mu_L = \mu_{max} \tag{3.3}$$

Expanding this expression to include its established power-law temperature dependence, the lattice scattering mobility is defined as

$$\mu_{i,L} = \mu_{max} \left(\frac{T}{300} \right)^{\theta_i} \tag{3.4}$$

where the subscript i specifies either electron or hole mobility and its corresponding model parameters [19].

Again examining Equation 3.2, the electron mobility due to donor scattering and the hole mobility due to acceptor scattering can be obtained by subtracting the lattice scattering mobility according to Matthiesen's rule. These majority impurity scattering mobilities, $\mu_{e,D}$ and $\mu_{h,A}$, are given by

$$\mu_{i,I}(N_I, c) = \mu_{i,N} \left(\frac{N_{ref,1}}{N_I} \right)^{\alpha_1} + \mu_{i,c} \left(\frac{c}{N_I} \right) \tag{3.5}$$

where (i, I) stands for (e, D) or (h, A), and c is the majority carrier concentration [18]. In the second term, the scaling factor of c/N_I accounts for impurity screening that begins to occur at high carrier concentrations. Including their temperature dependence, the parameters $\mu_{i,N}$ and $\mu_{i,c}$ are defined as [19]

$$\mu_{i,N} = \frac{\mu_{max}^2}{\mu_{max} - \mu_{min}} \left(\frac{T}{300} \right)^{3\alpha_1 - 1.5} \tag{3.6a}$$

$$\mu_{i,c} = \frac{\mu_{min}\mu_{max}}{\mu_{max} - \mu_{min}} \left(\frac{300}{T} \right)^{0.5} \tag{3.6b}$$

At both low temperatures and high carrier concentrations, it has been demonstrated that majority impurities scatter much more effectively than minority impurities. Using an analytical fit function $G(P_i)$ that addresses the dependencies on carrier concentration and temperature, the minority impurity scattering mobilities $\mu_{e,A}$ and $\mu_{h,D}$ are expressed as a simple ratio to the corresponding majority impurity scattering mobilities,

$$\mu_{e,A}(N_A, c) = \frac{\mu_{e,D}(N_D = N_A, c)}{G(P_e)} \tag{3.7a}$$

$$\mu_{h,D}(N_D, c) = \frac{\mu_{h,A}(N_A = N_D, c)}{G(P_h)} \tag{3.7b}$$

where $G(P_i)$ and $P_i(c, T)$ are as defined in [18].

Finally, the electron-hole scattering mobility can be expressed by considering electrons and holes as moving donors and moving acceptors, respectively. As such, an expression can be approached in the same way as for the minority impurity mobility. In this case, [18] proposes an analytical fit formula $F(P_i)$ that is used as a scaling factor on the majority impurity mobilities $\mu_{e,D}$ and $\mu_{h,A}$, with the resulting expressions for $\mu_{e,h}$ and $\mu_{h,e}$,

$$\mu_{e,h}(p, c) = F(P_e)\mu_{e,D}(N_D = p, c) \tag{3.8a}$$

$$\mu_{h,e}(n, c) = F(P_h)\mu_{h,A}(N_A = n, c) \tag{3.8b}$$

where $F(P_i)$ is as defined in [18].

Having developed analytical expressions for the mobility corresponding to each of the contributing scattering mechanisms, the overall mobility is computed following Matthiesen's rule,

$$\mu_i^{-1} = \mu_{i,L}^{-1} + \mu_{i,D}^{-1} + \mu_{i,A}^{-1} + \mu_{i,j}^{-1} \tag{3.9}$$

where $j = h$ if $i = e$ and $j = e$ if $i = h$. In [18], the foundations of this expression are further developed in order to account for effects such as weak screening and

the clustering of impurities at ultra-high concentrations. Furthermore, the majority impurity, minority impurity, and electron-hole scattering have all been derived for the case where there is only one type of scattering partner. Since in reality there are three types of potential scattering partners, it must be ensured that only truly two-body nearest-scatterers are counted among any of the possible scattering partners. The effective mobility from these three scattering mechanisms is given by

$$\mu_{i,D+A+j}(N_D, N_A, n, p) = \mu_{i,N} \frac{N_{i,\text{sc}}}{N_{i,\text{sc,eff}}} \left(\frac{N_{ref,1}}{N_{i,\text{sc}}} \right)^{\alpha_1} + \mu_{i,c} \left(\frac{n+p}{N_{i,\text{sc,eff}}} \right) \tag{3.10}$$

where $N_{i,\text{sc}}$ is equal to the sum of the majority impurity, minority impurity, and minority carrier concentrations. $N_{i,\text{sc,eff}}$ is equivalent to $N_{i,\text{sc}}$ except that the minority impurity concentration is scaled by $G(P_i)$ and the minority carrier concentration is divided by $F(P_i)$.

The strong temperature dependent nature of the lattice and majority impurity scattering mobilities is explicitly shown in Equations 3.4 and 3.6. For the majority impurity scattering mobility at low temperatures, $\mu_{i,N}$ will clearly dominate due to its direct power law dependence on temperature as opposed to the inverse power law dependence of $\mu_{i,c}$. The minority impurity and electron-hole scattering mobilities derive their temperature dependence both from their direct dependence on the majority impurity expression and from the parameter P_i within their respective mobility ratio functions $G(P_i)$ and $F(P_i)$. Assessing which of these scattering components drives the overall temperature dependence of the mobility at extremely low temperatures is an important step in evaluating and calibrating an accurate cryogenic mobility model. From Figure 3.1, it is clear that the lattice scattering mobility dominates the temperature dependence of the carrier mobility at lower doping concentrations and higher temperatures, whereas the combined majority/minority impurity and carrier scattering mobility increasingly dominates the temperature dependence for higher doping concentrations and lower temperatures. This provides a reasonable starting point for evaluating the temperature dependence of the lattice scattering mobility

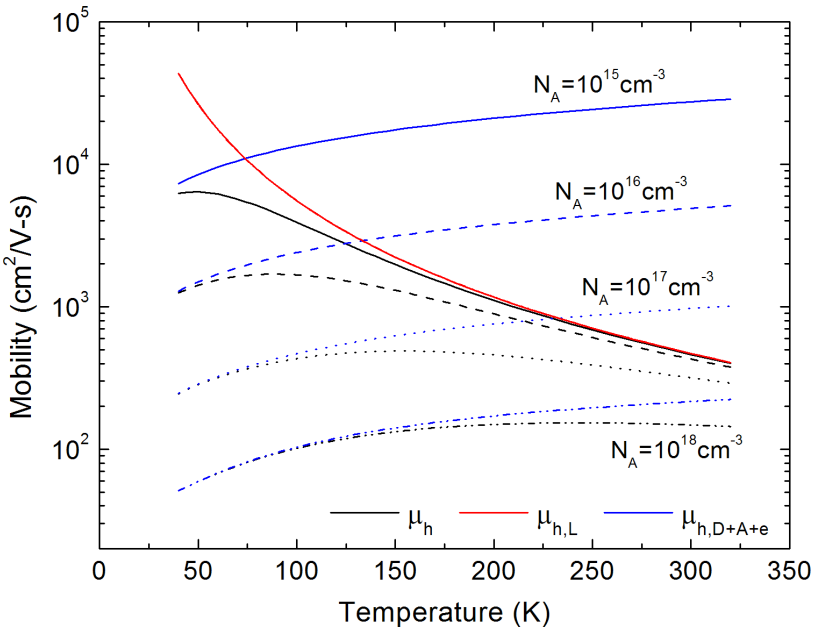


Figure 3.1: Comparison of the temperature dependence of scattering mechanisms used in Philips unified mobility model for various doping concentrations.

model against experimental resistivity measurements across temperature and doping concentration.

3.2.2 Incomplete Ionization

An accurate model for the incomplete ionization of dopants is necessary not only to meaningfully link experimental resistivity data to theoretical mobility values, but is in its own right a critical component of accurate low-temperature device models. Recently in two companion papers by Altermatt [2, 3] a parameterization of the density of states near the band edge of doped silicon is derived and subsequently applied to calculate the incomplete ionization of dopants.

Marked occupation of dopant states occurs when the Fermi level is located near the dopant level, leading to incomplete ionization of dopant atoms. This in turn results in a discernibly smaller free carrier density compared to the dopant density. Whereas incomplete ionization is commonly understood to be a significant low temperature effect, it is often neglected at higher temperatures, despite the demonstrated fact that

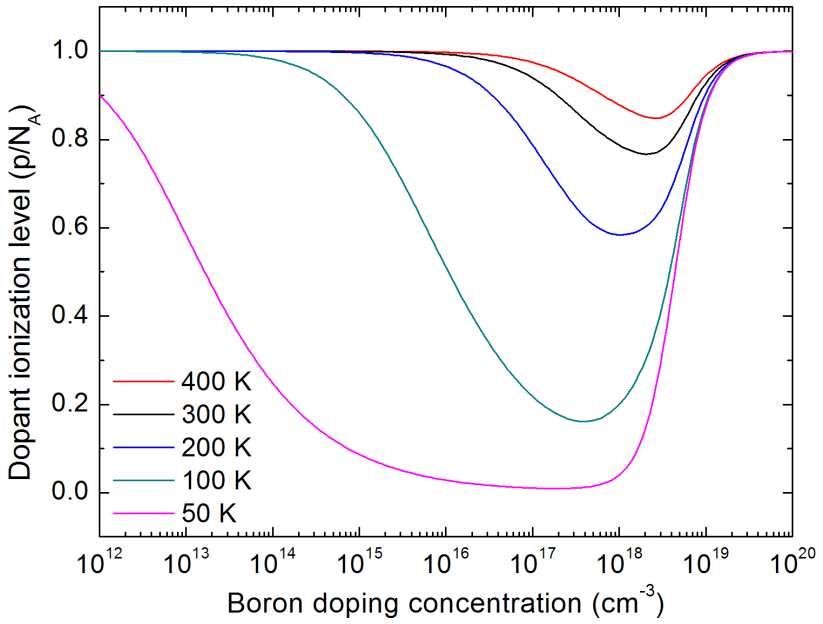


Figure 3.2: Ionization level as a function of boron doping concentration across a wide temperature range. These curves were calculated using the modified incomplete ionization model presented in [3, 45].

even at room temperature up to 25 % of dopant atoms may be nonionized for certain doping concentrations.

The model described in [2] is based on a parameterization of the density of states near the band edge of phosphorus-doped crystalline silicon that is derived from photoluminescence and conductance measurements, using a recently developed theory of band gap narrowing. In the model derivation, the dopant band is shown to only touch the conduction band at the Mott transition and to merge with the conduction band at considerably higher doping levels, agreeing with the experimentally demonstrated fact that at these high doping levels the dopants are completely ionized. This establishes that incomplete ionization at moderate temperatures is an important concern for doping levels from roughly $1 \times 10^{17}\,\mathrm{cm}^{-3}$ to $1 \times 10^{19}\,\mathrm{cm}^{-3}$, clearly shown in Figure 3.2, which demonstrates the temperature and doping dependence of dopant ionization.

In describing the incomplete ionization model, the case of n-type silicon will be

used. A suitable expression for the density of states near the conduction band can be given by the sum of the tail density of states $D_k(E)$ and a Gaussian density of states of the impurity band $D_{dop}(E)$. $D_{dop}(E)$ is multiplied by a doping dependent attenuation factor $b(N_{dop})$. A detailed derivation of $D_k(E)$ and $D_{dop}(E)$ is given in [2], including analytical expressions for each. Using these expressions, incomplete ionization can be computed by self-consistently solving the expressions for the concentrations of nonionized dopants, free carriers, and total electrically-active dopants. For n-type silicon, the following expressions are used:

$$N_{dop}^0 = \int_{-\infty}^{0} [D_{dop}(E) + D_k(E)] f_D(E, E_F) dE \tag{3.11a}$$

$$N_{dop}^+ = \int_{0}^{\infty} [D_{dop}(E) + D_k(E)] f(E, E_F) dE \tag{3.11b}$$

$$N_{dop} = N_{dop}^0 + N_{dop}^+ \tag{3.11c}$$

where N_{dop}^0 is the nonionized dopant density and N_{dop}^+ is the ionized dopant density. These quantities along with the Fermi level comprise the three unknowns of the above equations [45]. The equation parameters for phosphorus-doped silicon are given in [2], while the parameters for arsenic- and boron-doped silicon are given in [3].

Modifications to this rigorous ionization model are described in [3, 45], resulting in a simplified model that is suitable for implementation in a device simulator. Specifically, the tail density of states $D_k(E)$ is neglected, the Gaussian in the impurity density of states $D_{dop}(E)$ is replaced by a delta function, and the quasi-Fermi levels are expressed by densities using Maxwell-Boltzmann statistics. With these modifications, the expressions for the ionized dopant concentrations in n-type and p-type silicon are given as

$$\frac{N_{don}^+}{N_{don}} = 1 - \frac{b(N_{dop})n}{1 + ge^{-(E_{F_n} + E_{dop} - E_C)/kT}} = 1 - \frac{b(N_{dop})n}{n + gn_1} \tag{3.12a}$$

$$\frac{N_{acc}^-}{N_{acc}} = 1 - \frac{b(N_{dop})n}{1 + ge^{-(E_{dop} - E_{F_p} + E_V)/kT}} = 1 - \frac{b(N_{dop})p}{p + gp_1} \tag{3.12b}$$

where $b(N_{dop})$ represents the fraction of carriers in localized states, given by

$$b(N_{dop}) = (1 + (N_{dop}/N_b)^d)^{-1} \tag{3.13}$$

and the quasi-Fermi densities are expressed as

$$n_1 = N_C \exp(-E_{dop}/kT)\,, \quad p_1 = N_V \exp(-E_{dop}/kT) \tag{3.14}$$

with the screening-dependent binding energies E_{dop} modeled as

$$E_{dop} = \frac{E_{dop,0}}{1 + \left(\dfrac{N_{dop}}{N_{ref}}\right)^c} \tag{3.15}$$

This incomplete ionization model has been used for all relevant calculations in this paper, including the carrier lifetime modeling of the preceding chapter. It provides an accurate depiction of dopant ionization across doping concentration from the deep cryogenic through high temperature regimes. Together with the Philips mobility model, this model establishes a solid foundation for evaluating and calibrating an accurate resistance model based on the experimental measurements of resistivity that will be presented later in this chapter.

3.3 Experimental Conditions and Techniques

Having described the relevant theoretical models, experimental measurements of resistivity within a commercial SiGe BiCMOS process will be presented. In this section, the experimental conditions and measurement methods will be briefly reviewed.

All measurements were conducted within IBM's first generation (5HP/5AM) SiGe BiCMOS technology using an Agilent 4156 Semiconductor Parameter Analyzer to perform kelvin measurements on a variety of test structures. The kelvin measurement is a method to measure the actual resistance across an object without including any of the series resistance present in the measurement setup. Temperature dependent measurements of packaged test structures were carried out using a closed-cycle liquid-helium cryogenic test system capable of DC to 100 MHz operation from 4 K to 400 K.

Except for the p-type substrate, all of the doped regions that were measured are thin layers, most of which do not have constant doping profiles. Consequently, only the sheet resistances for these layers can be directly measured. For constant doping and a known layer thickness, the sheet resistance can simply be converted to the material resistivity, which can then be used to calibrate the theoretical models of mobility and incomplete ionization. However, in order to meaningfully link the theoretical models to the sheet resistance of a layer with variable doping, the models are first used to calculate the individual resistivities of a discretized doping profile; these resulting discretized sheet resistances are then used to calculate the total effective sheet resistance.

Measurements of p-type resistance include (1) the resistivity of the lightly doped substrate (constant doping on the order of $1 \times 10^{15}\,\mathrm{cm}^{-3}$), (2) the sheet resistance of the HBT base region, often referred to as r_{bi} (peak doping on the order of $1 \times 10^{18}\,\mathrm{cm}^{-3}$), and (3) the sheet resistance of the highly doped p+ diffusion layer used for PFET source and drain regions (peak doping greater than $1 \times 10^{20}\,\mathrm{cm}^{-3}$). Measurements of n-type resistance include the sheet resistances of (1) the lightly doped n-type epilayer from which the HBT collector is defined (constant doping near $5 \times 10^{15}\,\mathrm{cm}^{-3}$), (2) the collector region from the high-breakdown variation of the HBT (peak doping on the order of $5 \times 10^{16}\,\mathrm{cm}^{-3}$), and (3) the heavily doped subcollector region from the HBT (peak doping on the order of $1 \times 10^{20}\,\mathrm{cm}^{-3}$).

For the p-type substrate resistivity measurements, a test structure was designed according to the standard four-point-probe measurement. Four collinear 1.6 µm substrate contacts are equally spaced by 150 µm in order to avoid any influence on the overall measured resistance from the heavily doped region directly under the substrate contacts. The measured resistance R_{meas} of the substrate is expressed as the voltage across the inner two contacts divided by the current forced through the outer two

contacts. The resistivity of the substrate can then be determined by the expression

$$\rho_{sub} = 2\pi s R_{meas} \tag{3.16}$$

where s is the spacing between the contacts. The base sheet resistance r_{bi} was measured using a ring-dot structure, which is simply an emitter ring bounded by two inner and two outer base contacts. The cross-section of this structure can be visualized as similar to Figure 1.2, except that there are two separate base contacts on each side of the emitter. In order to avoid any Ge related effects, the particular structure that was measured did not include germanium grading in the base. The resistance for this structure is measured as the voltage across one pair of base contacts divided by the current driven through the alternate pair. The base sheet resistance is extracted by the expression

$$r_{bi} = \frac{R_{meas}}{2\pi \ln(R_1/R_0)} \tag{3.17}$$

where R_0 and R_1 are the inner and outer radii of the emitter ring, respectively. The remaining sheet resistances were measured using simple rectangular kelvin structures from which the sheet resistance is determined by scaling R_{meas} by a geometrical factor,

$$R_S = \frac{W}{L} \cdot R_{meas} \tag{3.18}$$

3.4 Experimental Results and Analysis

Measured data from the resistivity and sheet resistance measurements described above is presented in this section. First, an illustrative set of n-type resistivity measurements from [19] are displayed in Figure 3.3. As the doping concentration increases, the rise of the free carrier density outweighs the decrease in mobility and causes an overall decrease in resistivity. For lower doping concentrations, the decrease in resistivity with decreasing temperature is the result of a corresponding increase in mobility. In the low temperature regime, the effects of incomplete ionization can be seen as the resistivity increases due to a decreasing free carrier concentration. As the doping concentration

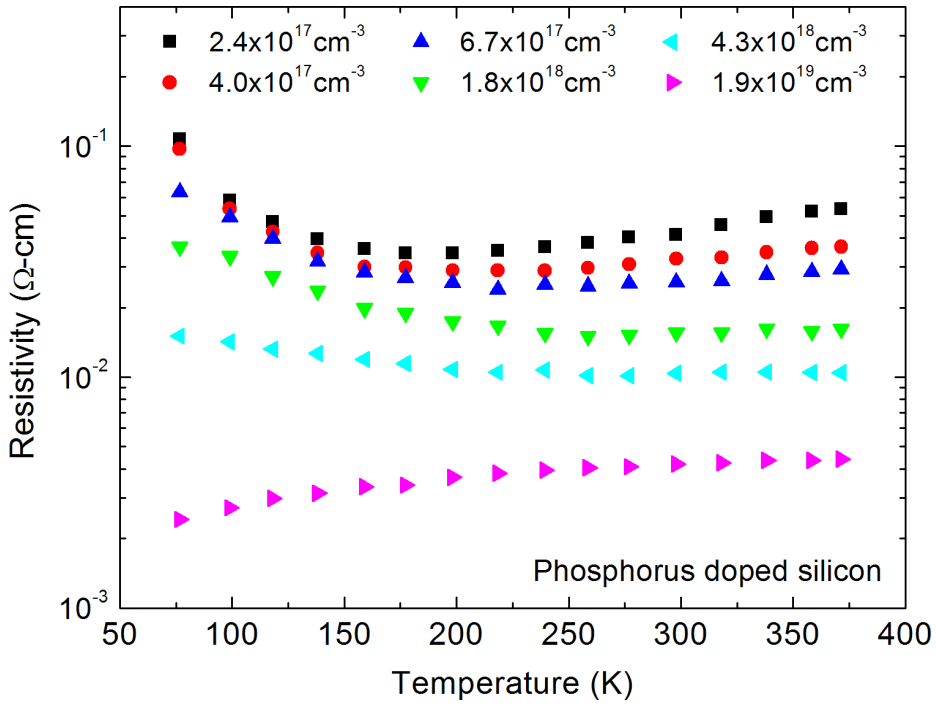


Figure 3.3: Resistivity of phosphorus doped silicon as a function of temperature for various doping concentrations (After [19]).

increases, the onset of incomplete ionization also rises with temperature, flattening out the temperature dependence of the resistivity at moderate temperatures. Finally, for very highly doped samples, the dopants remain completely ionized at all temperatures, leading to a resistivity temperature dependence that is entirely driven by the carrier mobility, which exhibits a moderate increase down through the cryogenic regime.

In Figure 3.4, the p-type temperature dependent resistance data is shown, including the substrate resistivity, base sheet resistance, and p+ diffusion sheet resistance. For the lightly doped substrate, a significant increase in resistivity is seen as the temperature decreases below 100 K; this can be attributed to the significant amount of incomplete ionization that is expected for a boron density of $9 \times 10^{14}\,\mathrm{cm}^{-3}$. The decrease in substrate resistivity from room temperature down to 100 K can be attributed solely to the expected increase in mobility, since the dopants are completely ionized in this temperature range. Unlike the substrate resistivity, the base sheet resistance exhibits clear signs of incomplete ionization even at temperatures above 200 K. The

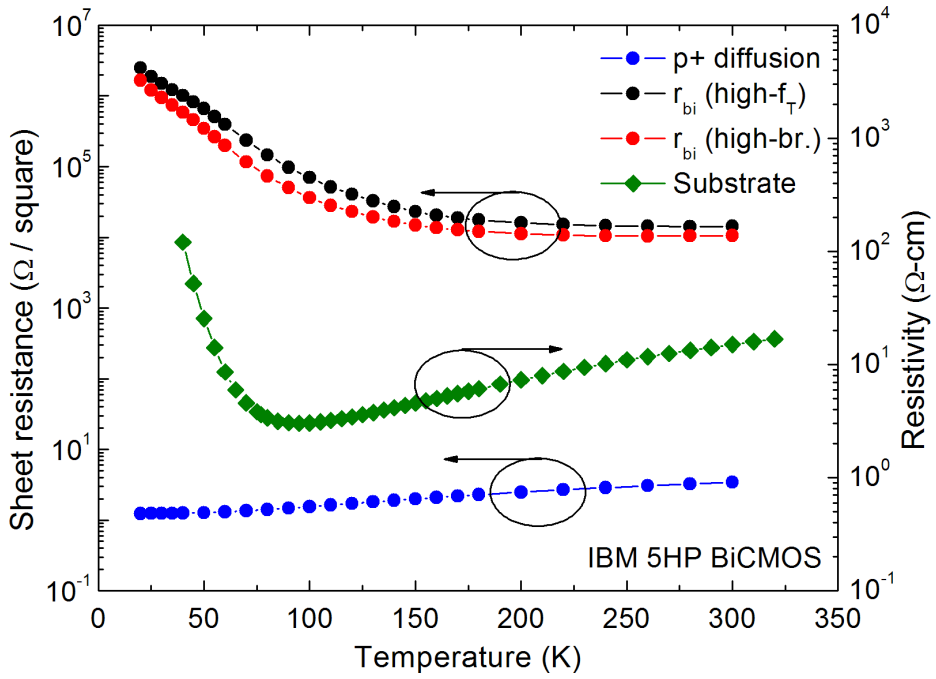


Figure 3.4: Resistivity and sheet resistance measurements for p-type silicon of various doping levels in IBM's 5HP process technology.

two base resistance curves are for separate ring-dot structures, one of which has the high-breakdown collector doping beneath the base layer, and one of which has the high-f_T collector doping. The curve shapes are virtually identical since the actual base profile is the same, but the higher collector doping within the high-f_T structure creates a higher base sheet resistance by compensating part of the base boron doping profile, effectively making the p-type base layer thinner. Reexamining Figure 3.2 for doping levels near $1 \times 10^{18}\,\mathrm{cm}^{-3}$, incomplete ionization is already in effect at room temperature and the level of ionization steadily decreases with decreasing temperature, albeit at a slower rate than for lower doping concentrations. This leads to an increase in r_{bi} with decreasing temperature, accelerating as the temperature decreases below 200 K. Finally, the p+ diffusion sheet resistance exhibits very little temperature dependence due to its extremely high doping concentration. Complete ionization of dopants holds across the entire temperature range, thus the slight decrease (less than 3x) in sheet resistance from 300 K to 20 K can be attributed to a corresponding

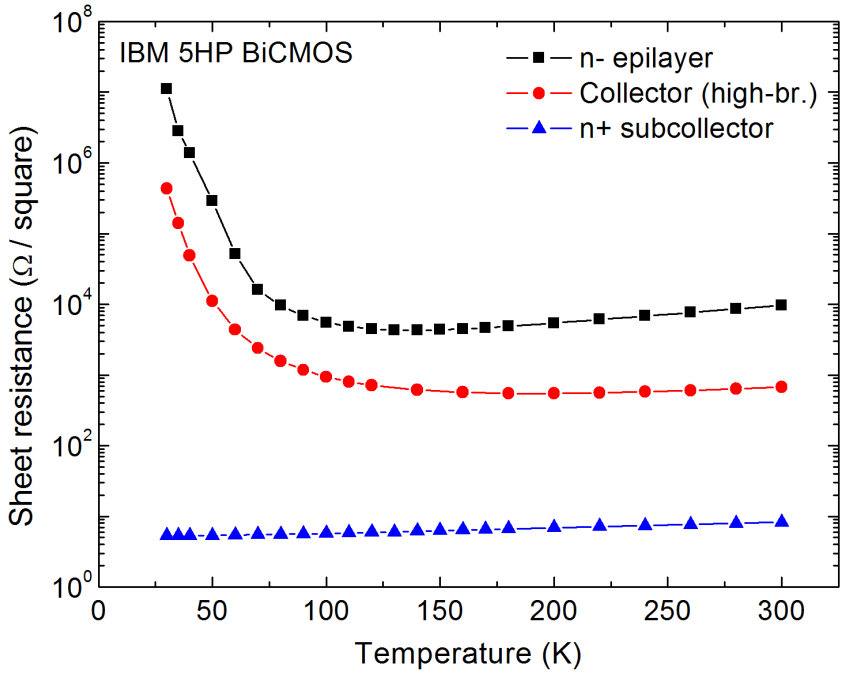


Figure 3.5: Sheet resistance measurements for n-type silicon of various doping levels in IBM's 5HP process technology.

increase in mobility.

Figure 3.5 shows the n-type temperature dependent resistance data, including the sheet resistances of the n- epilayer, HBT collector, and n+ HBT subcollector. For the n- epilayer, which has a relatively higher doping concentration than the substrate, a similar temperature dependence is recorded: the resistance displays a slow decreasing trend as the temperature decreases towards the cryogenic regime, then begins to quickly increase below 100 K as incomplete ionization takes effect. The weaker dependence at higher temperatures reflects the fact that the mobility dependence is also weaker due to the higher doping concentration (see Figure 3.1). Considering the collector layer, it is merely the n- epilayer after an ion implantation process step. The particular collector doping profile measured here is for the high-breakdown SiGe HBT, which has a lower peak doping concentration than that of the high-f_T SiGe HBT. Compared to the n- epilayer, the collector sheet resistance demonstrates a similar overall temperature dependence, with several enlightening differences: (1) the overall magnitude is lowered

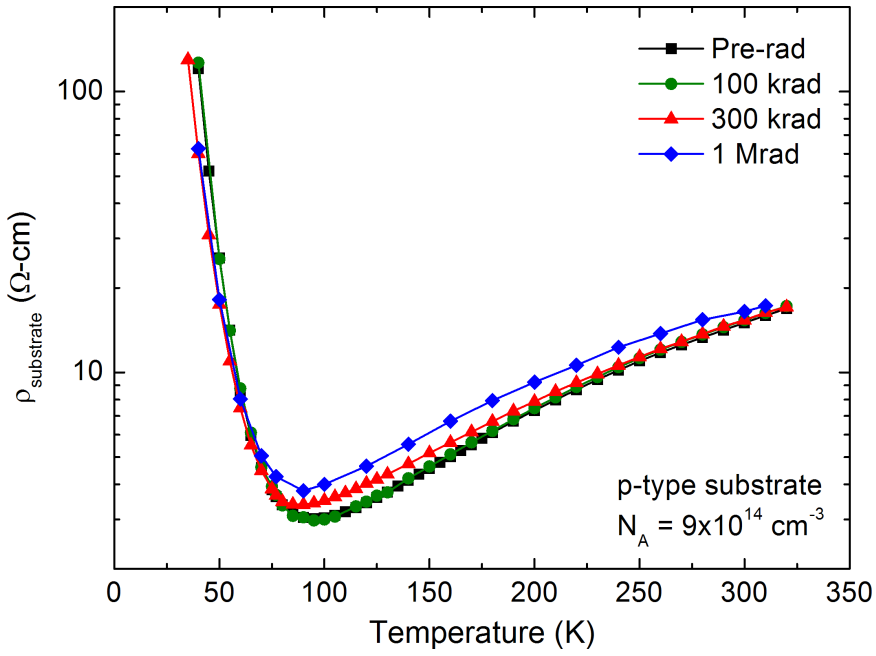


Figure 3.6: Proton radiation response of the substrate resistivity across temperature.

due to a higher carrier density, (2) the moderate temperature region is suppressed, reflecting a suppression of mobility from increased impurity scattering, and (3) the onset of incomplete ionization occurs at a higher temperature due to the higher doping concentration. Finally, the n+ subcollector resistance curve clearly indicates complete ionization across all temperatures. The decrease in resistance by a factor of 1.5 from 300 K to 20 K further indicates that the mobility only slightly increases with decreasing temperature. This result is in line with what we expect for extremely high doping concentrations.

Proton irradiation experiments were conducted at the Crocker Nuclear Laboratory at UC Davis in order to assess the effects of displacement and ionization damage on the substrate resistivity. Resistivity test structures were subjected at room temperature to 63 MeV proton irradiation up to a total accumulated dose of 1 Mrad (Si). Figure 3.6 shows the changes induced in the temperature dependent resistivity at accumulated doses of 100 krad, 300 krad, and 1 Mrad. The specific nature of the radiation induced changes to the resistivity is more clearly shown in Figure 3.7, in which the irradiated

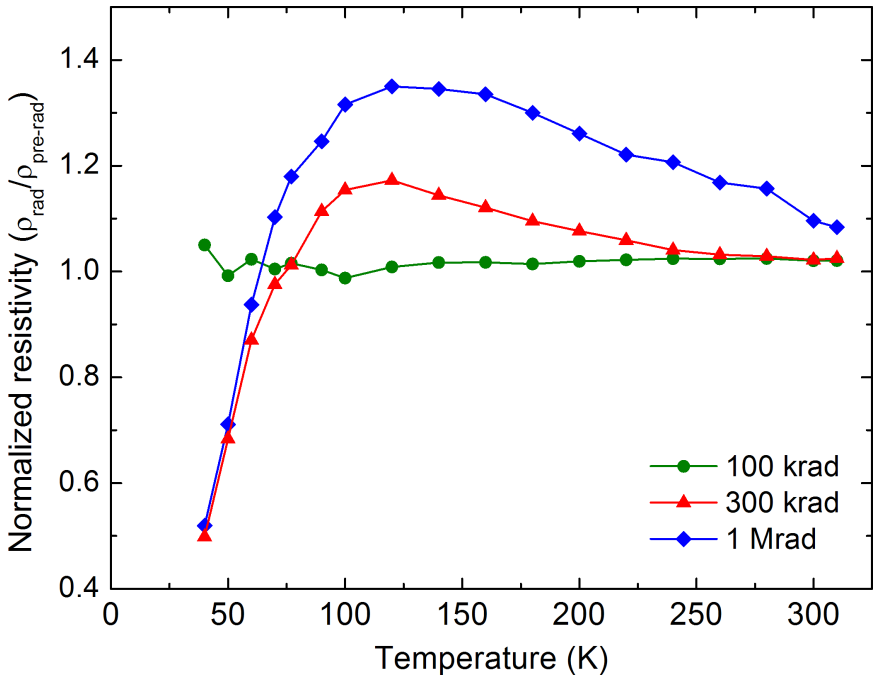


Figure 3.7: Relative change in substrate resistivity across temperature due to proton radiation.

resistivities are normalized to the pre-radiation data. At moderate temperatures, the increase in resistivity indicates that radiation induced displacement damage results in higher lattice scattering. Below 100 K, however, where impurity scattering dominates at this particular doping concentration, the resistivity actually decreases. One possible explanation for this decrease is that the radiation damage deactivates a fraction of boron dopant atoms. Dopant deactivation would cause a temperature-independent increase in resistivity due to a lower carrier concentration, along with decreased ionized impurity scattering that would be manifested as decreased resistance at low temperatures. Moreover, the lower carrier concentration would lead to a relatively lower degree of incomplete ionization, resulting in lower resistivity in the deep cryogenic temperature regime.

3.5 Modeling of Data

The temperature and doping dependent resistance data for n-type and p-type silicon can be used to develop calibrated theoretical models for integration into commercial TCAD software. In this section, this data will be considered in light of the mobility and incomplete ionization models presented in Section 3.2.

Although it is fairly easy to obtain an accurate model fit for a particular set of data by tweaking the model parameters, maintaining a high level of accuracy across a wide range of conditions is much more challenging. In attempting to reach mobility and ionization models that together produce accurate models of resistivity for the doping-dependent p-type and n-type resistance data presented here, the most reliable approach is to retain models that are physics-based and focused on material systems rather than particular technologies; this minimizes the reliance on assumptions that could potentially break down under conditions for which the models have not been experimentally evaluated. Models that are purely empirical or developed specifically for a particular technology often do not extend well to other technologies or physical conditions. Consequently, the approach in this paper has been to carefully calibrate the parameters of the Philips mobility model and the Altermatt ionization model, because both of these models were developed out of fundamental theory and aimed for silicon-based systems in general. The end goal of calibrating these models is to

Table 3.1: Parameters used in calibrated mobility model for arsenic-, phosphorus-, and boron-doped silicon.

Parameter	As	P	B
μ_{max} ($\mathrm{cm^2\,V^{-1}\,s^{-1}}$)	1417.0	1414.0	470.5
μ_{min} ($\mathrm{cm^2\,V^{-1}\,s^{-1}}$)	52.2	68.5	44.9
$N_{ref,1}$ ($\mathrm{cm^{-3}}$)	1.45×10^{17}	1.1×10^{17}	1.5×10^{17}
$N_{ref,I}$ ($\mathrm{cm^{-3}}$)	1×10^{22}	4×10^{20}	1×10^{22}
α_1	0.85	0.65	0.8
θ_i	1.72	1.72	1.82

achieve the best possible fits to all of the experimental data using a single set of model parameters. Accordingly, the modified parameters used for the calibrated Philips model and the calibrated Altermatt model are given in Tables 3.1 and 3.2, respectively.

In order to provide another reference for comparing the goodness of the model fits to the resistance data, an additional expression for the bulk mobility temperature dependence is considered, given in [1] as

$$\mu(T) = \mu_{max}\left(12.19\left(\frac{T}{300}\right)^4 - 33.56\left(\frac{T}{300}\right)^3 + 25.37\left(\frac{T}{300}\right)^2 - 4.07\left(\frac{T}{300}\right) + 1.08\right) \quad (3.19)$$

This expression, which was extracted from MOSFET-based measured data, is substituted in place of the power-law expression for the lattice scattering mobility in the full Philips model, while the remainder of the expressions for impurity and carrier scattering are left unchanged. Hereafter, this model will be referred to as the *UMd model*.

For resistivity and sheet resistance calculations, the following models were used: (1) the Philips mobility model combined with the Altermatt incomplete ionization model, with both models using the default parameters, (2) the UMd mobility model combined with the Altermatt model, with both models using the default parameters, and (3) the Philips mobility model combined with the Altermatt model, with both models using

Table 3.2: Parameters used in calibrated incomplete ionization model for arsenic-, phosphorus-, and boron-doped silicon.

Parameter	As	P	B
$E_{dop,0}$ (meV)	45.5	53.7	44.39
N_{ref} (cm^{-3})	7×10^{17}	3×10^{18}	8.5×10^{17}
c	0.8	1.5	1.4
N_b (cm^{-3})	6×10^{18}	9×10^{18}	4.5×10^{18}
d	1.3	1.8	2.4
g	1/2	1/2	1/4

the calibrated parameters presented in this paper. The third model is hereafter referred to as the *GT model*. In Figure 3.8, the measured temperature dependence of the substrate resistivity is shown alongside the calculated resistivities of the three model variations. Since lattice scattering dominates at moderate temperatures for a lightly doped sample, tuning θ_i proved crucial for fitting the data at these temperatures. The low temperature increase in resistivity due to incomplete ionization, combined with the increasing influence of impurity scattering, required simultaneous tuning of the impurity scattering temperature coefficient α_1 and the ionization model parameters.

Tuning of the ionization parameters necessarily has a significant impact on the calculated resistance for higher doping levels at which incomplete ionization is a concern across a wide temperature range. In this case, any tuning of the ionization parameters was reflected strongly in the calculated base resistance. Calculation of the base sheet resistance also required the doping variation across the base to be taken into account. Using the estimated boron profile taken from the SIMS profile (as in Figure 1.3), the base was discretized into very thin layers of constant doping. The resistivity and corresponding sheet resistance for each layer was calculated; subsequently, all of the individual sheet resistances were added together as parallel resistances in order to determine the total effective base sheet resistance. The modeled base sheet resistance is shown in Figure 3.9; although all three models give a reasonable fit for the base resistance, the GT model gives a slightly better fit, particularly down through 20 K. Moreover, with the same set of parameters, the GT model gives a significantly better fit to the lightly doped substrate resistivity data.

The modeled sheet resistance of the p+ diffusion layer is given in Figure 3.10. Although the shape of the calculated curves differs from that of the measured data, the relative error compared to the data remains fairly small down to 20 K. Since the dopants are completely ionized at this ultra-high doping level, any error in the models is due to inadequate modeling of ultra-high doping effects on mobility. Tuning the

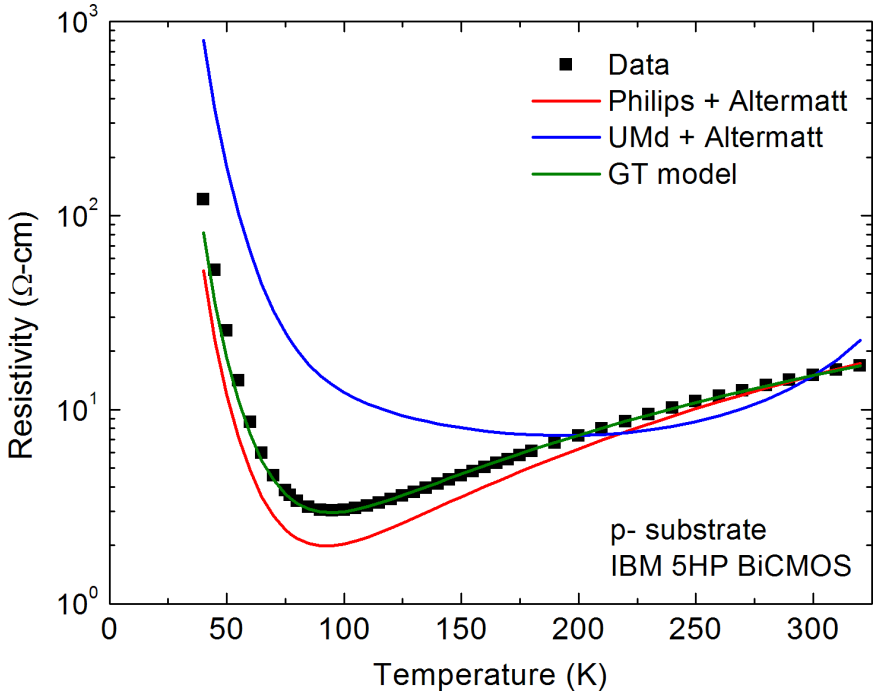


Figure 3.8: Comparison of model fits to substrate resistivity.

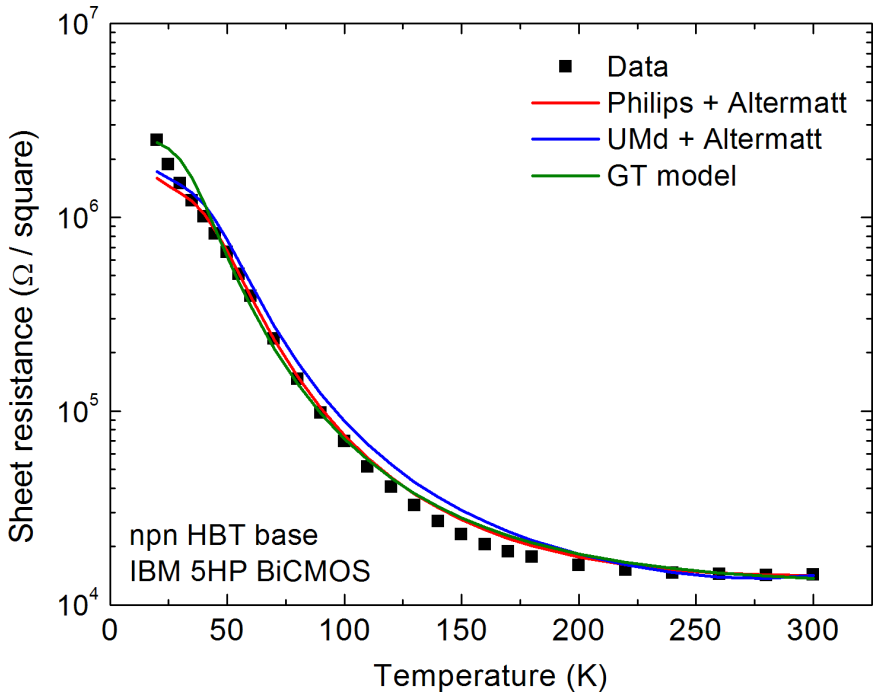


Figure 3.9: Comparison of model fits to npn HBT base sheet resistance.

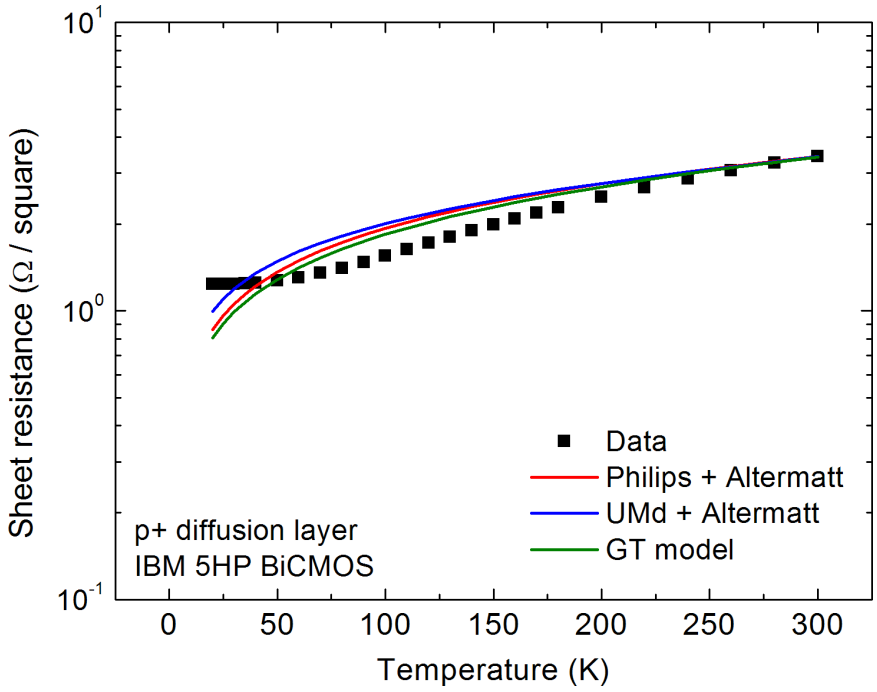


Figure 3.10: Comparison of model fits to p+ diffusion layer sheet resistance.

high-doping parameter $N_{ref,I}$ marginally improves the fit at moderate temperatures, but degrades it below 50 K. Moreover, the doping was assumed to be constant through the entire diffusion layer, while in reality a diffused layer has a highly variable doping profile. Although an effective doping concentration is a reasonable approximation since the majority of the current will conduct through the region surrounding the peak doping, effects such as freeze-out of the less highly doped fringe regions and variations in the mobility across doping will lead to some distortions of the measured resistance curve shape that are not accounted for by the models.

In Figure 3.11, the sheet resistance of the n- epilayer is shown alongside the calculated sheet resistance from each of the three models. The Philips and Altermatt models result in a very close fit to the data with the default parameter set; however, the fit is improved between 100 K and 250 K by tuning the lattice scattering temperature exponent θ_i. For these calculations, the doping level was assumed to be constant.

Likewise, the modeled sheet resistance of the HBT collector is shown in Figure

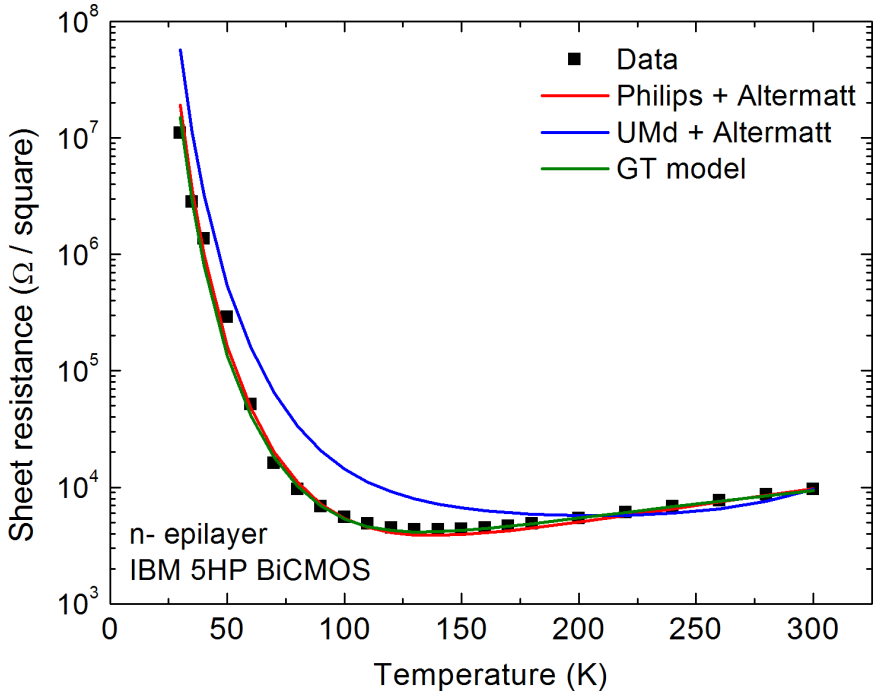


Figure 3.11: Comparison of model fits to n- epilayer sheet resistance.

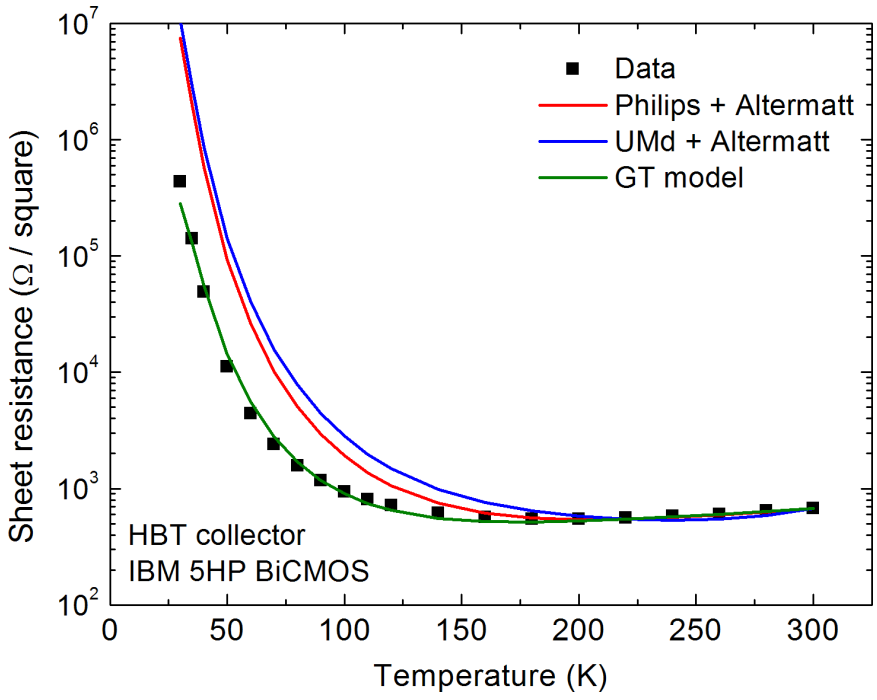


Figure 3.12: Comparison of model fits to npn HBT collector sheet resistance.

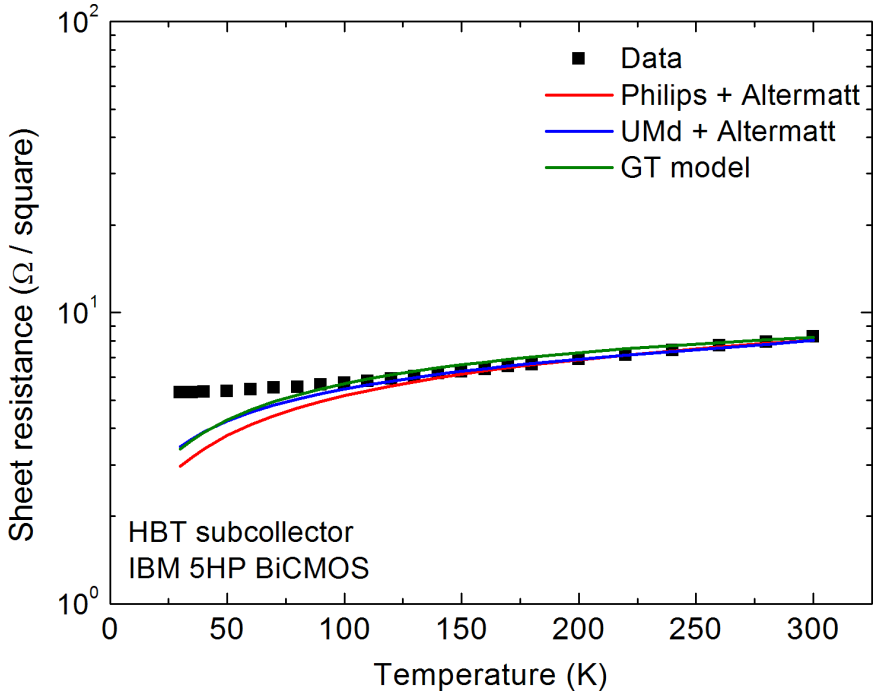


Figure 3.13: Comparison of model fits to npn HBT subcollector sheet resistance.

3.12. In this case, the doping level is higher than that of the epilayer, and as a result tuning of the incomplete ionization parameters is needed. Proper tuning results in an excellent fit to the data down to the lowest measured value at 30 K.

Finally, the modeled sheet resistance of the highly-doped HBT subcollector is shown alongside the measured data in Figure 3.13. Unlike the p+ diffusion layer, for which the doping profile was unknown, the subcollector profile was available for use in the same manner as the base profile; the profile was discretized in order to allow calculation of an effective sheet resistance that accounts for the variable doping of the subcollector layer. The three resistance models performed roughly the same, each diverging from the data at temperatures below 100 K and reaching values of roughly two to three times less than the data at 30 K. By tuning the ultra-high doping parameter $N_{ref,I}$, the GT model was able to perform slightly better than the default Philips model across temperature.

3.6 Summary

In this chapter, an overview was given of the approaches taken to model the temperature dependence of resistivity and mobility down to the deep cryogenic regime. State-of-the-art models for mobility and incomplete dopant ionization were addressed and subsequently applied to experimental resistance data taken from p-type and n-type silicon of various doping concentrations. Significant improvement in the accuracy of these models was demonstrated by simultaneously calibrating the model parameters to the measured data. In addition, the radiation response of the substrate resistivity was experimentally evaluated. Combined with the calibrated lifetime models presented in Chapter 2, the calibrated mobility and incomplete ionization models proposed here provide the necessary tools for accurate predictive TCAD modeling of charge carrier dynamics in the bulk regions of IBM's SiGe BiCMOS technology. Moreover, since advances in mobility and incomplete ionization modeling across doping concentration and temperature are applicable to nearly all aspects of SiGe HBT operation, they can be used to improve the overall accuracy of DC and AC device modeling.

CHAPTER IV

CONCLUSION

4.1 Conclusions

In the previous two chapters, three fundamental topics that drive carrier transport in silicon have been addressed: (1) carrier recombination, (2) carrier mobility, and (3) incomplete ionization of dopants. Combined together, accurate modeling of these physical phenomena is critical to any application in which the transient dynamics of charge carriers must be understood. Furthermore, mobility and dopant ionization are at the core of virtually all device performance metrics; this spans all types of devices, both HBT and MOSFET, and includes both DC and AC metrics. Despite the existence of well-defined recombination and mobility models for room-temperature operation, the extension of accurate models down to the deep cryogenic regime is a necessary step to developing viable technologies for use in applications such as space-based electronics.

The achievements presented in this thesis can be grouped into three categories. First, by way of extensive measurement, the behavior of minority carrier recombination lifetimes in both the substrate and n-well of a commercial SiGe BiCMOS technology have been characterized for the first time from high temperatures down through the cryogenic regime (Section 2.6). Also, the resistances of both p-type and n-type regions of various doping concentrations have been measured through the deep cryogenic temperature regime (Section 3.4). Second, the prevalent theoretical models for recombination lifetimes, carrier mobility, and incomplete ionization have been reviewed, with critical attention given to the low temperature limits of these models (Sections 2.3 and 3.2). Third, calibrated theoretical models for recombination lifetimes,

carrier mobility, and incomplete ionization have been presented (Sections 2.7 and 3.5). The accuracy of the calibrated lifetime model was demonstrated against both the injection and temperature dependent data, whereas the accuracy of the calibrated mobility and ionization models was shown against both doping and temperature dependent resistance data. Moreover, the effects of proton irradiation on the substrate recombination lifetime were closely accounted for using the recombination model presented here.

4.2 TCAD Integration

Having calibrated the SRH model to fit the temperature and injection dependence of carrier lifetime data, a desirable application would be to implement it within standard TCAD software. Sentaurus Device [55], part of the Sentaurus software suite offered by Synopsys, implements the SRH recombination model similar in essence to that which was presented in Section 2.3, taking as inputs the room temperature lifetime (*taumax*), lifetime temperature dependence, and trap energy level. In order to implement our model, the trap parameters given in Table 2.2 can be easily translated and inserted as SRH model parameters within the silicon parameter file, assuming that a single trap level is dominant. Sentaurus Device allows either a exponential or power law dependence of lifetime on temperature. As such, since the temperature dependence of the capture cross-section is simply the inverse of the lifetime temperature dependence, it can be inverted and directly input to the software. The trap energy level can be specified directly, and the trap density can be accounted for by calibrating the room temperature lifetime to match that which is predicted by the model.

Similarly, properly calibrated Philips mobility and Altermatt incomplete ionization models demonstrated accurate fits to the measured data in Section 3.5. In order to capitalize on these models, the must be integrable with standard TCAD tools. As Sentaurus Device has just been shown to be compatible with a calibrated carrier

lifetime model, it would be advantageous if it were also compatible with calibrated mobility and incomplete ionization models. Fortunately, Sentaurus Device directly implements the Philips mobility model, and in the same manner in which the trap parameters can be specified, the mobility model parameters can also be specified in the silicon parameter file.

Unlike the Philips model, however, Sentaurus Device does not currently implement the Altermatt incomplete ionization model. This can be circumvented, since Sentaurus Device allows the definition of custom models through the "Physical Model Interface"; the only limitation is that these models must follow a pre-defined form. As described in [45], Equation 3.12 may be transformed into the following form so that it matches the pre-defined functionality built into the software:

$$\frac{N_{don}^{+}}{N_{don}} = \frac{1}{1 + g_D n/n_1} \tag{4.1a}$$

$$\frac{N_{acc}^{-}}{N_{acc}} = \frac{1}{1 + g_A n/n_1} \tag{4.1b}$$

where the effective degeneracy factors g_D and g_A are given as

$$g_D(T, N_{don}) = \frac{b}{g + (1-b)N_{don}/n_1} \tag{4.2a}$$

$$g_A(T, N_{acc}) = \frac{b}{1 + (1-b)N_{acc}/p_1} \tag{4.2b}$$

The use of these effective degeneracy factors ensures a smooth transition to complete ionization at high doping concentrations, unlike constant degeneracy factors, which lead to a monotonically increasing degree of incomplete ionization with increasing doping concentrations.

With the successful implementation of these three calibrated models, accurate predictive TCAD simulations are now possible, enabling a host of different simulation applications; several potential applications for these models will be addressed in the following section.

4.3 Future Research

As was described in Section 1.4.2, single-event-effects in SiGe digital circuits are a critical concern that must be addressed in order for SiGe technology to be a viable option for space-based electronics. Although a number of device and circuit level techniques have been developed in order to mitigate either the amount of transient charge collected or the effect of current spikes on circuit operation, the need for further study remains, especially as SiGe based technologies evolve. Among the many approaches to studying SEE, it is critical to first understand the fundamental nature of charge collection at the terminals of a single HBT after an ion strike. This is done by 2D and 3D TCAD simulations that capture the transient charge transport in the device and the bulk regions surrounding it. With this type of simulation setup, charge collection can be investigated as a function of ion species, ion energy, strike angle, and strike location. Clearly, however, the fundamental physical models that govern charge transport behavior will drastically affect the behavior of the simulated charge collection. Using the calibrated models in this thesis, future research could be directed towards implementing accurate predictive TCAD simulations of charge collection. After confirming the accuracy of these simulations against empirical measurements of charge collection, these simulations will provide additional leverage for understanding and developing SEE immune SiGe devices and circuits.

Further research can also be conducted into refining the mobility and ionization models presented in Section 3.5. In particular, thoroughly capturing the behavior of mobility and incomplete ionization requires characterization of the temperature dependence of resistivity at many more doping concentrations. In this paper, the modeling was limited by inadequate knowledge of the specific doping concentrations and layer thicknesses. Additional measurements under which these parameters are well known would decrease the uncertainty and inaccuracy of the calibrated models. Moreover, direct measurements of mobility and/or dopant ionization across doping

concentration and temperature would allow each model to be calibrated individually without the uncertainty that is introduced by inaccuracies in the other model.

Finally, an interesting application of the calibrated models proposed in this thesis would be to extend them to compact modeling of SiGe HBTs. For example, the temperature dependence of the base, collector, and subcollector sheet resistances has been carefully modeled in this paper. These are critical parameters that determine key performance metrics such as f_{MAX}. Although TCAD simulations are critical to designing semiconductor devices and understanding their fundamental operation, compact models are arguably more important, because they are the foundation of all circuit design. Consequently, it would be extremely beneficial if advances in accurate device modeling at the TCAD level could be extended to the corresponding compact models. Furthermore, it may even be easier to calibrate components of a compact model, since for example the temperature parameters can be customized for each individual terminal resistance rather than for the material system in general.

Many other applications exist in addition to these. Nearly all TCAD simulations within this specific SiGe BiCMOS technology will benefit from improved modeling of recombination, mobility, and incomplete ionization—even more so since these models are calibrated to experimental measurements from the same technology that is being simulated. As a result, even though in some cases these calibrated models will only serve to strengthen theoretical foundations, the work presented here has the potential for a subtle, but widespread influence within the field of SiGe technology as a whole.

REFERENCES

[1] Akturk, A., Allnutt, J., Dilli, Z., Goldsman, N., and Peckerar, M., "Device modeling at cryogenic temperatures: Effects of incomplete ionization," *IEEE Transactions on Electron Devices*, vol. 54, no. 11, pp. 2984–2990, 2007.

[2] Altermatt, P. P., Schenk, A., and Heiser, G., "A simulation model for the density of states and for incomplete ionization in crystalline silicon. I. Establishing the model in Si:P," *Journal of Applied Physics*, vol. 100, Dec. 2006.

[3] Altermatt, P. P., Schenk, A., and Heiser, G., "A simulation model for the density of states and for incomplete ionization in crystalline silicon. II. Investigation of Si:As and Si:B and usage in device simulation," *Journal of Applied Physics*, vol. 100, Dec. 2006.

[4] Babcock, J. A., Cressler, J. D., Vempati, L. S., Clark, S. D., Jaeger, R. C., and Harame, D. L., "Ionizing radiation tolerance of high performance SiGe HBTs grown by UHV/CVD," *IEEE Transactions on Nuclear Science*, vol. 42, pp. 1558–1566, 1995.

[5] Basore, P. and Hansen, B. R., "Microwave-detected photoconductance decay," *Conference Record of the Twenty First IEEE Photovoltaic Specialists Conference, 1990*, pp. 374–379, 1990.

[6] Bean, J. C., Sheng, T. T., Feldman, L. C., Fiory, A. T., and Lynch, R. T., "Pseudomorphic growth of Ge_xSi_{1-x} on silicon by molecular beam epitaxy," *Applied Physics Letters*, vol. 44, pp. 102–104, 1984.

[7] Birkholz, J. E., Bothe, K., Macdonald, D., and Schmidt, J., "Electronic properties of iron-boron pairs in crystalline silicon by temperature- and injection-level-dependent lifetime measurements," *Journal of Applied Physics*, vol. 97, p. 103708, 2005.

[8] Bleichner, H., Jonsson, P., Keskitalo, N., and Nordlander, E., "Temperature and injection dependence of the Shockley-Read-Hall lifetime in electron-irradiated n-type silicon," *Journal of Applied Physics*, vol. 79, pp. 9142–9148, June 1996.

[9] Claeys, C. and Simoen, E., *Radiation Effects in Advanced Semiconductor Materials and Devices*. Berlin: Springer-Verlag, 2002.

[10] Cressler, J. D., ed., *The Silicon Heterostructure Handbook: Materials, Fabrication, Devices, Circuits, and Applications of SiGe and Si Strained-Layer Epitaxy*. CRC Press, 2006.

[11] Cressler, J. D. and Niu, G., *Silicon-Germanium Heterojunction Bipolar Transistors*. Boston: Artech House, 2003.

[12] Fischer, A., Osten, H.-J., and Richter, H., "An equilibrium model for buried SiGe strained layers," *Solid State Electronics*, vol. 44, pp. 869–873, 2000.

[13] Goebel, H. and Hoffmann, K., "Full dynamic power diode model including temperature behavior for use in circuit simulators," *International Symposium on Power Semiconductor Devices and ICs*, pp. 130–135, 1992.

[14] Green, M. A., "Minority carrier lifetimes using compensated differential open circuit voltage decay," *Solid-State Electronics*, vol. 26, no. 11, pp. 1117–1122, 1983.

[15] Iyer, S. S., Patton, G. L., Delage, S. S., Tiwari, S., and Stork, J. M. C., "Silicon-germanium base heterojunction bipolar transistors by molecular beam epitaxy," *International Electron Devices Meeting*, vol. 33, pp. 874–876, 1987.

[16] Keskitalo, N., Jonsson, P., Nordgren, K., Bleichner, H., and Nordlander, E., "Temperature and injection dependence of the Shockley-Read-Hall lifetime in electron-irradiated p-type silicon," *Journal of Applied Physics*, vol. 83, pp. 4206–4212, Apr. 1998.

[17] Klaassen, D. B. M., "A unified mobility model for device simulation," *International Electron Devices Meeting Technical Digest*, pp. 357–360, Dec. 1990.

[18] Klaassen, D. B. M., "A unified mobility model for device simulation–I. Model equations and concentration dependence," *Solid-State Electronics*, vol. 35, no. 7, pp. 953–959, 1992.

[19] Klaassen, D. B. M., "A unified mobility model for device simulation–II. Temperature dependence of carrier mobility and lifetime," *Solid-State Electronics*, vol. 35, no. 7, pp. 961–967, 1992.

[20] Krithivasan, R., Niu, G., Cressler, J. D., Currie, S. M., Fritz, K. E., Reed, R. A., Marshall, P. W., Riggs, P. A., Randall, B. A., and Gilbert, B., "An SEU hardening approach for high-speed SiGe HBT digital logic," *IEEE Transactions on Nuclear Science*, vol. 50, pp. 1849–1854, Dec. 2003.

[21] Lang, D. V., "Deep level transient spectroscopy: A new method to characterize traps in semiconductors," *Journal of Applied Physics*, vol. 45, pp. 3023–3032, 1974.

[22] Larin, F., *Radiation Effects in Semiconductor Devices*. New York: John Wiley and Sons, 1968.

[23] Lu, Y., Cressler, J. D., Krithivasan, R., Li, Y., Reed, R. A., Marshall, P. W., Polar, C., Freeman, G., and Ahlgren, D., "Proton tolerance of a third generation, 0.12 μm 185 GHz SiGe HBT technology," *IEEE Transactions on Nuclear Science*, vol. 50, pp. 1811–1815, Dec. 2003.

[24] Macdonald, D., "Impact of nickel contamination on carrier recombination in n-and p-type crystalline silicon wafers," *Applied Physics A*, vol. 81, pp. 1619–1625, 2005.

[25] Macdonald, D. and Cuévas, A., "Validity of simplified Shockley-Read-Hall statistics for modeling carrier lifetimes in crystalline silicon," *Physical Review B*, vol. 67, no. 7, p. 75203, 2003.

[26] Macdonald, D., Cuevas, A., Rein, S., Lichtner, P., and Glunz, S. W., "Temperature- and injection-dependent lifetime spectroscopy of copper-related defects in silicon," *3rd World Conference on Photovoltaic Energy Conversion*, pp. 87–90, May 2003.

[27] Macdonald, D., Kerr, M., and Cuevas, A., "Boron-related minority-carrier trapping centers in p-type silicon," *Applied Physics Letters*, vol. 75, no. 11, pp. 1571–1573, 1999.

[28] Marshall, C. and Marshall, P., "Proton effects and test issues for satellite designers - Part B: Displacement effects," *IEEE Nuclear Space and Radiation Effects Conference*, 1999.

[29] Marshall, P. and Marshall, C., "Proton effects and test issues for satellite designers - Part A: Ionization effects," *IEEE Nuclear Space and Radiation Effects Conference*, 1999.

[30] Marshall, P. W., Carts, M. A., Campbell, A., McMorrow, D., Buchner, S., Stewart, R., Randall, B., Gilbert, B., and Reed, R. A., "Single event effects in circuit hardened SiGe HBTs at gigabit per second data rates," *IEEE Transactions on Nuclear Science*, vol. 47, pp. 2669–2674, Dec. 2000.

[31] Meyerson, B. S., "Low-temperature silicon epitaxy by ultrahigh vacuum/chemical vapor deposition," *Applied Physics Letters*, vol. 48, pp. 797–799, Mar. 1986.

[32] Mukherjee, T. S., Sutton, A. K., Kornegay, K. T., Krithivasan, R., Cressler, J. D., Niu, G., and Marshall, P. W., "A novel circuit-level SEU hardening technique for high-speed SiGe HBT digital logic circuits," *IEEE Transactions on Nuclear Science*, vol. 54, pp. 2086–2091, Dec. 2007.

[33] NASA Space Environments and Effects (SEE) Program, *NASA SEE TWG - Ionosphere and Thermosphere*. Available at http://see.msfc.nasa.gov/pf/pf.htm.

[34] Niu, G., Krithivasan, R., Cressler, J. D., Marshall, P., Marshall, C., Reed, R., and Harame, D., "Modeling of single event effects in circuit-hardened high-speed SiGe HBT logic," *IEEE Transactions on Nuclear Science*, vol. 48, pp. 2126–2134, 2001.

[35] Patton, G. L., Harame, D. L., Stork, J. M. C., Meyerson, B. S., Scilla, G. J., and Ganin, E., "Graded-SiGe-base, poly-emitter heterojunction bipolar transistors," *IEEE Electron Device Letters*, vol. 10, pp. 534–536, Dec. 1989.

[36] Pierret, R. F., *Advanced Semiconductor Fundamentals*. Reading, MA: Addison-Wesley, 1989.

[37] Pierret, R. F., *Semiconductor Device Fundamentals*. Reading, MA: Addison-Wesley, 1996.

[38] Reed, R. A., Marshall, P. W., Ainspan, H., Marshall, C. J., Kim, H. S., Cressler, J. D., Niu, G., and LaBel, K. A., "Single event upset test results on an IBM prescalar fabricated in IBM's 5HP germanium doped silicon process," *Proc. IEEE Nuclear Space Radiation Effects Conference Data Workshop*, pp. 172–176, 2001.

[39] Rein, S. and Glunz, S. W., "Electronic properties of the metastable defect in boron-doped czochralski silicon: Unambiguous determination by advanced lifetime spectroscopy," *Applied Physics Letters*, vol. 82, pp. 1054–1056, Feb. 2003.

[40] Rein, S. and Glunz, S. W., "Electronic properties of interstitial iron and iron-boron pairs determined by means of advanced lifetime spectroscopy," *Journal of Applied Physics*, vol. 98, p. 103711, 2005.

[41] Rein, S., *Lifetime Spectroscopy: A Method of Defect Characterization in Silicon for Photovoltaic Applications*. Springer, 2005.

[42] Roldán, J., Ansley, W. E., Cressler, J. D., Clark, S., and Nguyen-Ngoc, D., "Neutron radiation tolerance of advanced UHV/CVD SiGe HBTs," *IEEE Transactions on Nuclear Science*, vol. 44, pp. 1965–1973, 1997.

[43] Roldán, J., Niu, G., Ansley, W. E., Cressler, J. D., Clark, S. D., and Ahlgren, D. C., "An investigation of the spatial location of proton-induced traps in SiGe HBTs," *IEEE Transactions on Nuclear Science*, vol. 45, pp. 2424–2430, 1998.

[44] Rosenits, P., Roth, T., Glunz, S. W., and Belijakowa, S., "Determining the defect parameters of the deep aluminum-related defect center in silicon," *Applied Physics Letters*, vol. 91, p. 122109, 2007.

[45] Schenk, A., Altermatt, P. P., and Schmithusen, B., "Physical model of incomplete ionization for silicon device simulation," *Simulation of Semiconductor Processes and Devices, International Conference on*, pp. 51–54, 2006.

[46] Schmidt, J., "Temperature- and injection-dependent lifetime spectroscopy for the characterization of defect centers in semiconductors," *Applied Physics Letters*, vol. 82, pp. 2178–2180, Mar. 2003.

[47] Schmidt, J., Bothe, K., and Hezel, R., "Oxygen-related minority-carrier trapping centers in p-type czochralski silicon," *Applied Physics Letters*, vol. 80, no. 23, pp. 4395–4397, 2002.

[48] Schmidt, J. and Sinton, R., "Lifetime spectroscopy for defect characterization: Systematic analysis of the possibilities and restrictions," *Journal of Applied Physics*, vol. 91, pp. 2059–2070, Feb. 2002.

[49] Schmidt, J. and Sinton, R., "Defect characterization by temperature and injection-dependent lifetime spectroscopy," *3rd World Conference on Photovoltaic Energy Conversion*, pp. 947–950, May 2003.

[50] Sinton, R. A. and Cuevas, A., "Contactless determination of current-voltage characteristics and minority-carrier lifetimes in semiconductors from quasi-steady-state photoconductance data," *Applied Physics Letters*, vol. 69, no. 17, pp. 2510–2512, 1996.

[51] Stassinopolous, E. G., Brucker, G. J., Nakamura, D. W., Staufer, C. A., Gee, G. B., and Barth, J. L., "Solar flare proton evaluation at geostationary orbits for engineering applications," *IEEE Transactions on Nuclear Science*, vol. 43, pp. 369–382, 1996.

[52] Sutton, A. K., Haugerud, B., Lu, Y., Kuo, W.-M., Cressler, J., Marshall, P. W., Reed, R. A., Rieh, J.-S., Freeman, G., and Ahlgren, D., "Proton tolerance of fourth-generation 350 GHz UHV/CVD SiGe HBTs," *IEEE Transactions on Nuclear Science*, vol. 51, pp. 3736–3742, Dec. 2004.

[53] Sutton, A. K., Krithivasan, R., Cressler, J. D., Pellish, J. A., Reed, R. A., Marshall, P. W., Varadharajperumal, M., Niu, G., and Vizkelethy, G., "An evaluation of transistor-layout RHBD techniques for SEE mitigation in SiGe HBTs," *IEEE Transactions on Nuclear Science*, vol. 54, pp. 2044–2052, Dec. 2007.

[54] Sutton, A. K., Moen, K., Cressler, J. D., Carts, M. A., Marshall, P. W., Pellish, J. A., Ramachandran, V., Reed, R. A., Alles, M. L., and Niu, G., "Proton-induced SEU in SiGe digital logic at cryogenic temperatures," *Solid-State Electronics*, vol. 52, pp. 1652–1659, 2008.

[55] Synopsys, Inc., *Sentaurus Device User Guide, Version A-2008.09*, Sept. 2008.